VOL. 17 No. 9, pp. 233-268 MARCH 1956

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
THE PHILIPS INDUSTRIES

EDITED BY THE RESEARCH LABORATORY OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN, EINDHOVEN, NETHERLANDS

SIMPLE THEORY OF THE JUNCTION TRANSISTOR

by F. H. STIELTJES and L. J. TUMMERS. 537.311.33:621.375.4

In the future development of the electronic industry, the transistor is undoubtedly destined to play a more and more prominent role. It is quite conceivable that the use of this crystal amplifier will come to be as frequently discussed in this Review as the electronic tube. It therefore seems desirable to devote a series of articles to the basic theory of the transistor and its behaviour in circuits. This first article deals with the simple theory of the transistor, applicable for small current densities and low frequencies (audio frequencies).

Introduction

A transistor is a crystal amplifier which can be used to amplify electrical signals. It is manufactured from semi-conducting material, for which, so far, germanium has chiefly been used. The "junction transistor" derives its name and its properties from the so-called *P-N* junctions between the different layers of which it is composed [1]). The properties of these junctions have been discussed in an earlier article in this Review [2]) (further referred to as I), in which the rectifying action of a *P-N* junction was explained. So far as they are necessary for an understanding of the transistor mechanism, we may recapitulate the ideas given in I.

P and N regions in germanium; P-N junction

In germanium the electrical conductivity is due to two kinds of charge carriers: the "conduction electrons" and the "conduction holes", subsequently referred to simply as "electrons" and "holes". The holes manifest themselves as positive, mobile charge carriers. In the absence of outside influences, at every point in the crystal, the product of the concentration of holes and the concentration of electrons is equal to a constant (which increases rapidly with rising temperature), characteristic for germanium. The presence of traces of certain elements in the crystal lattice makes the concentration of holes very high and, in view of the constancy of the product, that of the electrons low; this gives us *P*-germanium. Other elements raise the concentration of the electrons and thus lower that of the holes; one then speaks of *N*-germanium. The charge carriers in the higher and lower concentrations are called the majority and minority charge carriers respectively; the minority carriers are thus electrons in the *P*-region and holes in the *N*-region.

In the simple theory of the *P-N* junction and also of the transistor, we may assume that if the foreign chemical elements present are homogeneously distributed in their respective regions, every volume element is electrically neutral [3]). Together with the condition that the product of the concentrations of holes and electrons be constant, the neutrality condition determines the values of these two concentrations in homogeneous regions of the crystal in the absence of external influences. The values of the concentrations thus fixed are known as the "equilibrium concentrations".

P and *N* regions may be present in the same

[1]) The idea of the junction transistor was originated by Shockley, who, after setting out the theory of the *P-N* junction, showed how the junction transistor could be constructed and predicted its properties. See Bell Syst. tech. J. **28**, 435-489, 1949.

[2]) J. C. van Vessem, Theory and construction of germanium diodes, Philips tech. Rev. **16**, 213-224, 1954/55 (No. 8).

[3]) Strictly speaking, space charges can occur at the extreme outer edges of a homogeneous crystal region. The distances over which these space charge regions extend, however, are always so small that by comparison with the other distances involved in the theory of the transistor, they may be neglected.

crystal. At the junction between P and N regions large space charge densities may occur. The space charge region, which is very thin (see also note [3])), and which one can visualize as being bounded by two more or less sharply defined planes *1* and *2*, is called the "P-N junction" or "barrier". *Fig. 1a*

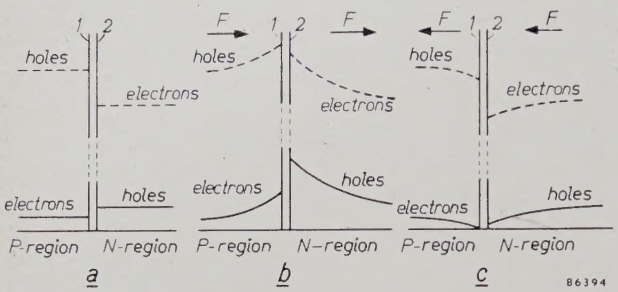


Fig. 1. Concentrations of minority charge carriers (full lines) and of majority charge carriers (broken lines), as functions of position on either side of a P-N junction in a germanium crystal. The concentration of majority charge carriers in practice is between 10^4 and 10^6 times that of the minority charge carriers. *1* and *2* denote the boundary planes of the junction.
a) No voltage is applied to the crystal; everywhere outside the junction the concentrations assume uniform equilibrium values. By the addition of suitable impurity elements in the proper proportions, both in the P-region and the N-region, the equilibrium concentrations can be varied within wide limits (on condition that the product of minorities and majorities in any one region remains constant).
b) A voltage is applied to the crystal which raises the potential on the P-region with respect to the N-region. The electric field F, whose direction is shown by arrows, drives majority charge carriers (holes in the P-region and electrons in the N-region) towards the P-N junction.
c) The polarity of the voltage is opposite to that in (*b*). The field now drives majority charge carriers away from the junction.

represents the equilibrium concentrations of majority and minority charge carriers in the P and N regions, supposed homogeneous, on opposite sides of such a barrier. How the concentrations vary within the barrier is not of importance for the simple theory of the transistor. A typical value of the concentration of the minority charge carriers is, for example, 10^{11} per cm^3; that of the majority charge carriers is then 5×10^{15} per cm^3 (since the constant product is about 5×10^{26} at room temperature).

In the following discussion, we shall consider principally the minority charge carriers. We shall see that the properties of the P-N junction and of the transistor are largely determined by the behaviour of the minority charge carriers, although the majority carriers are indispensable to the physical processes to be described. Furthermore it will be seen that the currents of minority carriers (or, briefly, minority currents) can be directly calculated and that conclusions can then easily be drawn about the majority currents.

The concentration of the minority charge carriers (or, briefly, minority concentration) will be denoted in general by g and its equilibrium value by G. Where necessary a suffix will indicate to which region the concentration applies.

Concentration changes under an applied voltage

If a voltage is applied between a P and an N region, then deviations from the equilibrium concentrations occur in the neighbourhood of the junction (fig. 1*b* and 1*c*). If the portion of the potential difference which is developed between the two boundary planes *1* and *2* of the P-N junction be V (V is considered positive if the potential of the P-region is raised with respect to that of the N-region), then the minority concentrations in these boundary planes are given by:

$$g = G\, e^{Vq/kT} \quad \ldots \ldots \quad (1)$$

where $-q$ is the charge of the electron, k is Boltzmann's constant and T the absolute temperature [4]). For each of the two boundary planes the appropriate equilibrium concentration must be substituted for G.

For a given P-N junction the minority concentrations in the boundary planes are thus fixed by the external voltage across the junction. It can be shown that when a voltage is applied, the neutrality condition remains valid for homogeneous regions of the crystal outside the barrier layer [5]), but that the condition that the product of the minority and majority concentrations be constant then ceases to apply. The neutrality condition means that outside the barrier layer, at every point in a homogeneous crystal region where the minority concentration is altered, the majority concentration must also be altered by an equal amount.

Diffusion currents of minority charge carriers; forward and reverse direction

As explained at length in I, pairs of electrons and holes continually appear and disappear: so-called generations and recombinations. Generations will be dominant where the concentrations lie below the equilibrium values, and recombinations will dominate in the reverse case. Concentrations differing from the equilibrium values will be set up at the barrier planes when a voltage is applied across the junction. Some distance from the junction, the equilibrium concentration will be maintained by

---

[4]) In accordance with the increasing practice in transistor literature, the symbol q is used for the electronic charge in place of the conventional e. In article I, q/kT was introduced as an unspecified constant a; V corresponds to $-\Delta\varphi$ in article I, while (1) corresponds to equations (16*a*) and (16*b*) in that article.

[5]) This neutrality condition was used in I to deduce equation (1).

generation and recombination: thus a concentration gradient is set up which is steepest near the junction and trails off into the crystal. Whenever a gradient exists in the concentration of particles that take part in the thermal motion, the thermal agitation sends, on an average, more particles in the direction from high to low concentration than vice versa. This is the well-known phenomenon of diffusion. The magnitude of the diffusion current density at any point is (apart from the sign) proportional to the concentration gradient at that point; the proportionality factor D is called the "diffusion constant".

The charge carriers are subjected to the action of the electric field as well as to that of diffusion: the resulting current is a superposition of the field current and the diffusion current. Where the majority current is concerned, the field and diffusion work in opposite directions; it can be seen in fig. 1*b* that the field drives holes to the right, while the concentration gradient in the *P*-region drives them to the left. The field, however, which set the whole process in motion, retains the upper hand. For the minority carriers, the field and diffusion reinforce each other. *The minority current resulting from the electric field is, however, everywhere negligible with respect to the total current (majority plus minority currents).* The reason for this lies in the great difference in the concentrations: since the field current is proportional to the concentration, a field which produces a reasonable current of majority carriers will result in no appreciable minority current. When the minority current makes an appreciable contribution to the total current, this contribution must be derived almost entirely from diffusion. *Where the minority current is concerned, therefore, we need to consider only the diffusion current.* Thus the flow of holes which, in fig. 1*b*, crosses the *P-N* junction into the *N*-region is almost purely diffusion current. This current is maintained by a continuous supply of holes from the *P*-region. In fig. 1*c*, the diffusion current of holes in the *N*-region is directed towards the *P-N* junction. Those holes which reach the junction are carried off into the *P*-region by the electric field against the diffusion action. The *electron* currents in fig. 1*b* and 1*c* are brought about in an analogous manner.

That the minority charge carriers are almost exclusively moved by diffusion, can be seen in a more quantitative way as follows.

The current of holes I^+ consists of a portion I_F^+, supplied by the field, and a portion I_d^+, supplied by diffusion. The same applies to the electron current I^-, which consists of I_F^- and I_d^-. Suppose that in the *N*-region, the field supplies a fraction A of the minority current (i.e. hole current). Then:

$$I_F^+ = A\, I^+,$$

and thus

$$I_d^+ = (1-A)\, I^+.$$

The field will give rise to a majority current of electrons, which is greater than the field hole current by a certain factor B, thus:

$$I_F^- = B\, I_F^+ = BAI^+.$$

We are here considering only cases where the majority concentration is everywhere very much greater than the minority concentration. The factor B is then a very large number: for example of the order of magnitude of 10^5.

Since the concentration of electrons and holes follow parallel curves, the concentration gradients for the two sorts of charge carriers are the same in any cross-section. The diffusion constant for electrons is about twice as great as that for holes, and the diffusion current of electrons is thus a factor $M \approx 2$ greater than that of holes. As a result of the difference in sign of the charge, the currents, regarded electrically, are in opposite directions:

$$I_d^- = -M\, I_d^+ = -M\,(1-A)\, I^+.$$

The total current is

$$I = I^+ + I^- = I^+ + I_F^- + I_d^- = \{1 + BA - M(1-A)\}\, I^+,$$

so that

$$\frac{I^+}{I} = \frac{1}{1-M+(B+M)\,A}. \quad \ldots\ldots \quad (2)$$

At some distance from the *P-N* junction, where the concentration gradient is rather small, the current of minority charge carriers may be as much as 1% field current, i.e. $A = 10^{-2}$. Then (taking $B = 10^5$ and $M = 2$) I^+ is about one thousandth of the total current and the whole minority current is thus negligible. This case is of no interest. Everywhere where I^+ is not negligible compared to the total current, however, A is very small. It can easily be calculated that $A \approx 10^{-3}$ when $I^+/I = 1\%$, and A decreases as I^+/I increases. If the minority current is appreciable, then it is almost entirely a diffusion current.

At a great distance from the *P-N* junction, where the equilibrium concentrations are substantially the equilibrium values, the concentration gradient is very small. Diffusion is then quite negligible and the minority current is therefore purely a field current. However, the whole minority current is then negligible, in accordance with the conclusions drawn from (2).

The fact that the behaviour of the minority carriers is determined entirely by diffusion is the reason why their behaviour is much simpler to analyse than that of the majority carriers. In article I (p. 221) it was deduced for the one-dimensional case, with which we are also concerned here, that the minority concentration approaches the equilibrium value as an exponential function of distance x. The distance L in which the deviation from the equilibrium value changes by a factor e is the

so-called diffusion length [6]). The fact that we are here dealing with an exponential function gives us the following relationship (see *fig. 2*):

$$\frac{\mathrm{d}g(x)}{\mathrm{d}x} = \frac{g(x) - G}{L}. \quad . \quad . \quad . \quad . \quad . \quad (3)$$

For the minority current [7]), which is actually a diffusion current (I_d),

$$I_\mathrm{d} = q\,D\,\frac{\mathrm{d}g(x)}{\mathrm{d}x},$$

or, in combination with (3):

$$I_\mathrm{d} = q\,D\,\frac{g(x) - G}{L}. \quad . \quad . \quad . \quad . \quad . \quad (4)$$

In these formulae no account is taken of the sign of $\mathrm{d}g(x)/\mathrm{d}x$ and of I_d; the sign of I_d can be decided by a glance at the figure. The diffusion constant for holes D_p, or that for electrons D_n, must be substituted for D according as the minority current consists of holes or electrons. D_n is about twice as large as D_p. The value of L is strongly dependent on lattice defects in the crystal and on the impurities present, and its value can vary widely as between the various homogeneous regions of the crystal. In the following discussions we shall, where necessary, indicate by a suffix to L, the region to which this quantity relates [8]).

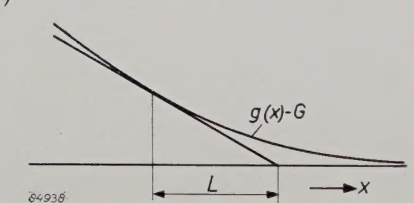


Fig. 2. The difference $g(x) - G$ between the concentration $g(x)$ of minority charge carriers and their equilibrium concentration G follows, in the one-dimensional case, an exponential curve. The sub-tangent of such a function is constant and represents the distance in which the function changes its value by a factor e (here, this is the diffusion length L) Thus for every value of x, ignoring the sign, $\mathrm{d}g(x)\mathrm{d}x = \{g(x) - G\}/L$.

Since $g(x)$, for the boundary planes, is given by (1), we can express the minority currents in these planes in terms of the voltage across the *P-N* junction:

$$I_\mathrm{d} = q\,\frac{DG}{L}\,(\mathrm{e}^{Vq/kT} - 1) \quad . \quad . \quad . \quad . \quad . \quad (5)$$

The total current is, of course, the sum of hole and electron currents calculated at the same place. Consider the cross-section *2* at the boundary plane of the junction (fig. 1). Here the holes form the minority current, the magnitude of which is given by (5). The electron current at *2*, i.e. the majority current, is unknown. However, we now make use of the fact that, because the *P-N* junction is very thin, the effect of generation and recombination within the junction is negligible: hence the electron current across *2*, where it is the majority current, must be equal to that across plane *1*, where it is the minority current, and is given by calculating the latter from (5).

Fig. 3a and *b* show the concentrations of the minority charge carriers at a *P-N* junction and their relation to the diffusion lengths, for both voltage directions. From these diagrams it is simple to deduce the minority currents at the junction, and thus also the value of the total current. Fig. 3*c* and *d* show how the hole current and the electron current change continuously on going from the *P*-region to the *N*-region and vice versa, so that their roles as majority and minority currents are interchanged.

As V becomes increasingly negative, it is seen from (5) that the diffusion current approaches a saturation level. Also from fig. 3*b* it is immediately clear that saturation must occur: the minority concentrations cannot fall below zero. Since at room temperature $kT/q \approx {}^1/_{40}$ volt, the saturation value of the current has been reached long before the potential drop across the *P-N* junction is -1 volt. The negative sign indicates that the potential is in the reverse direction. If V is positive, the current increases rapidly with the voltage; this is the forward direction.

The following remarks will serve to illustrate that the electrical conductivities of the adjoining P and N regions have a very unexpected effect on the current across the junction. As we have seen, the total current can be evaluated by adding the minority diffusion currents at the junction as given (5). by At the junction, each type of carrier therefore contributes an amount proportional to its minority equilibrium concentration G to the total current (the latter, of course, is the same throughout the crystal). Now, the minority equilibrium concentrations G will be smaller in a crystal of high conductivity since then the concentration of majority carriers is high (product of equilibrium concentrations is constant). Hence we find the apparent paradox that for a given voltage across the junction, the current is higher the worse the conductivities of the P and N regions on either side of it. This again illustrates that it is not the electrical conductivity but the diffusion which determines the current.

---

6) Diffusion-recombination length would perhaps be a more appropriate name since L is the average distance which minority charge carriers move in the x-direction by diffusion, before they disappear by recombination.

7) Strictly speaking, not the current, but the current density is of importance here. One can imagine simply that we are here dealing with a *P-N* junction of unit cross-sectional area.

8) Once more we deviate from the notation in I. In I the suffices to the diffusion lengths L_n and L_p mentioned on p. 221 refer to the type of charge carrier.

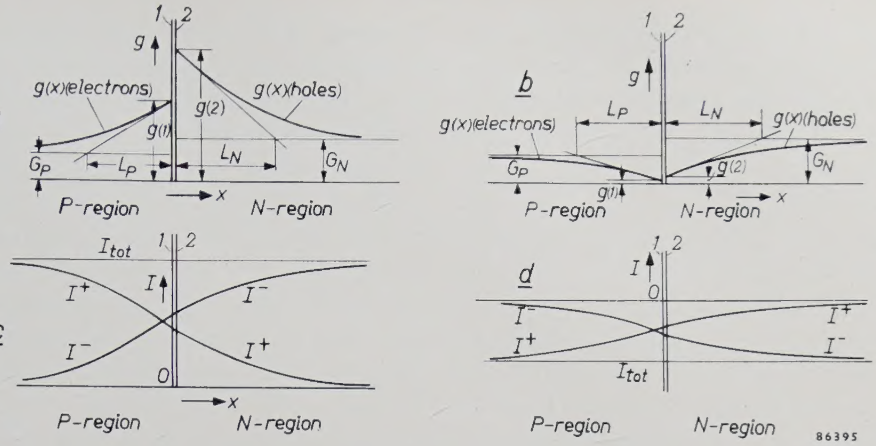


Fig. 3. Variation of the concentration g of the minority charge carriers (electrons in the P-region and holes in the N-region) at the P-N junction. The concentration gradients occurring at the boundary planes *1* and *2* of the junction can be read off the figure, after which the minority currents themselves can be written down directly. L_P and L_N are the diffusion lengths in the P and N regions respectively.
a) The voltage across the junction is in positive direction of x; the field drives holes to the right (same case as in fig. 1*b*).
b) The voltage is of the same magnitude, but negative; the field drives holes to the left (same case as in fig. 1*c*).
c) Hole current I^+ and electron current I^- for the P-N junction in situation (*a*). In the direction of increasing x, the hole current gradually gives way to the electron current, so that the total current remains constant.
d) Hole current and electron current in situation (*b*). Both currents are now negative and, furthermore, much smaller than in situation (*a*): the voltage is in the reverse direction, whereas in (*a*) it is in the forward direction.

Action of the junction transistor

Basis of amplification by a transistor

A transistor consists of a single crystal of germanium in which a P-N junction facing in the x-direction is followed by an N-P junction. There are thus two P-regions, separated by an N-region, known as the base. Each region is fitted with an electrode (*fig. 4*).

Before investigating the details of the action of the transistor, we shall first summarize its essentials.

At one boundary plane of the base (cross-section *2*, see *fig. 5*) the minority concentration (holes) is fixed at a value $g(2)$ lying above the equilibrium value by applying a voltage across the P-N junction in the *forward* direction. In the other end-plane (section 3) the concentration of holes is fixed at zero by a sufficiently large voltage across the N-P junction in the *reverse* direction. A large concentration gradient in the x-direction now exists in the base.

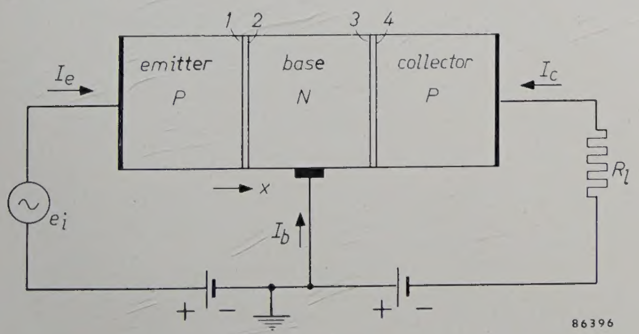


Fig. 4. Schematic representation of a P-N-P-transistor. Between the P-regions to the left and right, respectively designated "emitter" and "collector", is the "base" of N-germanium. I_e, I_c and I_b are respectively the emitter, collector and base currents. Currents in the direction of the arrows are considered positive. *1*, *2*, *3* and *4* mark the boundary planes of the P-N and N-P junctions. e_i A.C. input source; R_l resistance in the collector circuit, across which the amplified voltage appears.

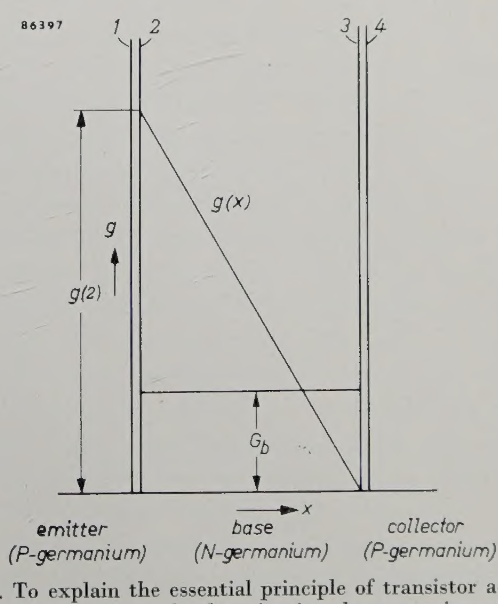


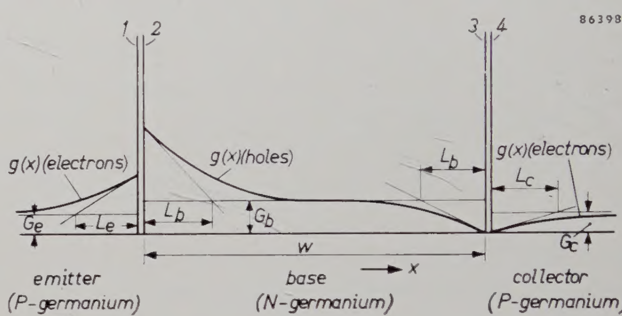
Fig. 5. To explain the essential principle of transistor action, the concentration of only the minority charge carriers $g(x)$ in the base of a junction transistor is given as function of position x. G_b represents the equilibrium concentration of the minority charge carriers in the base.

Since the current of minority charge carriers is a diffusion current, the concentration gradient results in a current of holes in the x-direction. This flow of holes is maintained by a continuous supply of holes from the P-region on the left; this P-region is therefore called the "emitter". It is often said that the emitter "injects" holes into the base. The holes disappear over the N-P junction into the P-region to the right; this P-region is therefore known as the "collector". The current of holes through the base thus makes a contribution to the current through the collector circuit. Since the emitter voltage is in the forward direction with respect to the base, a small variation of this voltage gives a large change in the concentration of holes at section *2* (see eq. 1). This also causes a change in the concentration gradient and thus in the flow of holes in the base and across the collector junction. The result is a change of current in the collector circuit and thus in the voltage drop over a load resistance R_l which is included in the circuit (see fig. 4). This is accompanied by a change in the potential of the collector with respect to the base, but since this potential is in the reverse direction, the collector voltage can vary quite considerably before the concentration of holes at section *3* deviates appreciably from zero. It is in fact possible to make R_l so large that a small change in the voltage between emitter and base reappears, amplified, across R_l. In the circuit shown (fig. 4), the transistor thus gives voltage amplification.

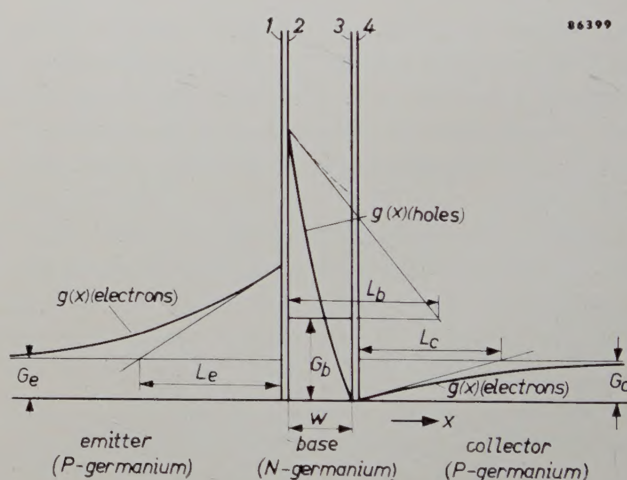
Closer examination of the situation in a transistor

In the above explanation of the amplification produced by a transistor, no mention was made of the electrons, nor of generation and recombination. We shall now fill in these gaps. In *fig. 6a* it is assumed that the base thickness w is considerably greater than the diffusion length of holes in the base ($w \gg L_b$). The variation of the minority concentrations in emitter, base and collector is shown: in the emitter and collector this is the concentration of electrons; in the base, that of holes. Again, the minority concentration at the emitter junction is raised to a level above the equilibrium value by a voltage in the forward direction, while at the collector junction it is held almost at zero by a voltage in the reverse direction. The currents across the two junctions are independent of each other; they are simply the currents across single P-N junctions as discussed earlier. The difference between the currents across the emitter and collector junctions is supplied via the base contact. Because a change of voltage between emitter and base has no effect on the concentration gradient at the collector junction, there is, in this case, no transistor action.

In fig. 6*b* it is assumed that the base thickness is small with respect to the diffusion length ($w \ll L_b$), as should be the case in a good transistor. In contrast to the case shown in fig. 6*a*, the concentration gradient for holes in the base near the collector (section 3) is now no longer equal to G_b/L_b as would follow from (eq. 3), but greater, because the concentration of holes $g(2)$ at section *2* exerts an influence (of magnitude dependent on the thickness w of the base). The current of holes crossing the collector junction will not, therefore, have the normal saturation value corresponding to the given G_b and L_b, but will be larger. A change in $g(2)$, occasioned by a change of voltage between emitter and base, will now certainly influence the flow of holes across



a



b

Fig. 6. Minority concentrations $g(x)$ in the emitter and in the collector (electrons) and in the base (holes) of a P-N-P transistor, as functions of position x.
a) The base thickness w is large with respect to the diffusion length L_b in the base. The concentration curve is determined by the generation and recombination processes, just as in separate P-N junctions. The concentration gradient of holes at section *3* is G_b/L_b and is not influenced by the hole concentration in section *2*. There is no transistor action.
b) The base thickness w is small compared to L_b. The hole concentration in section *3* is now affected by the concentration at section *2*, and transistor action occurs. In emitter and collector, the concentration curves remain the same as in (*a*). The broken line shows the hole concentration curve that would occur in the base if the N-P junction was much further away from the P-N junction.

the collector junction. The *electron* current in the collector is not altered because the concentration gradient for electrons in the collector is not altered. A variation of the hole current across the collector junction thus means that the same variation occurs in the *total* current through the collector circuit, for there is no flow of holes to or from the base contact (the concentration gradient is directed substantially along the x-axis and the hole current follows this gradient).

If the thickness w of the base is very small compared with the diffusion length, generation and recombination in the base can be neglected. The hole current, given by $qD\mathrm{d}g/\mathrm{d}x$, is then constant throughout the base, and since q and D are constants, the same applies to the concentration gradient $\mathrm{d}g/\mathrm{d}x$, which becomes $g(2)/w$. This gives the situation shown in fig. 5. The bend in the concentration curve, as shown in fig. 6*b*, is caused by the fact that generation and recombination do have some effect.

At the emitter, the current change which accompanies the voltage change is supplied partly by electrons and partly by holes. Since only holes contribute to the current change in the collector circuit, the transistor, in the circuit shown in fig. 4, gives a current amplification of less than 1. However, the product of current amplification and voltage amplification can, by careful design of the transistor and suitable choice of R_l and of the D.C. bias voltages of the emitter and collector circuits, be made much larger than unity, so that power amplification occurs.

Emitter efficiency γ, base efficiency β and current amplification factor a

For a small increase in the emitter voltage V_e (considered positive when in the forward direction), the current across the emitter junction increases by an amount ΔI_e, consisting of a hole portion ΔI_e^+ and an electron portion ΔI_e^-. Since only ΔI_e^+ is effective in producing a change of collector current, the transistor works better the larger the fraction ΔI_e^+ makes up of the total current change ΔI_e. For constant collector voltage V_c, this fraction is called the emitter efficiency γ, thus

$$\gamma = \left(\frac{\Delta I_e^+}{\Delta I_e}\right)_{V_c}. \quad \dots \dots (6)$$

If generation and recombination in the base cannot be neglected, it should be noted that recombination will be dominant because the average hole concentration in the base lies above the equilibrium value (fig. 6*b*). Only a fraction β of ΔI_e^+ reaches the collector to bring about an increase ΔI_c^+ in the collector current. This fraction, again measured at constant collector voltage, is called the base or transport efficiency:

$$\left(\frac{\Delta I_c^+}{\Delta I_e^+}\right)_{V_c} = -\beta. \quad \dots \dots (7)$$

(The negative sign comes from the convention given in fig. 4 of calling currents positive if they are directed towards the crystal.) The electron current across the collector, as we have said, is constant because the concentration gradient for the electrons in the collector remains unchanged. This applies in the case considered here ($V_c =$ constant) exactly. Thus:

$$\Delta I_c = \Delta I_c^+ = -\beta \Delta I_e^+ = -\beta\gamma \Delta I_e.$$

This relation fixes the current amplification which occurs at constant collector voltage. The positive product $\beta\gamma$ is called the current amplification factor a; if we now change over to differentials, this may be written:

$$a = -\left(\frac{\partial I_c}{\partial I_e}\right)_{V_c}. \quad \dots \dots (8)$$

Also,

$$a = \beta\gamma. \quad \dots \dots (9)$$

From the definition of β and γ it follows that the current amplification factor $a < 1$. In practice V_c will not be constant and the real current amplification will be smaller than a. Efforts are made to make β and γ independently as close to unity as possible in order to make a as large as possible.

Approximate calculation of γ

Neglecting generation and recombination in the base, it is easy to calculate γ. It follows from (1) that the ratio of the minority concentrations on opposite sides of a *P-N* junction is equal to that of their equilibrium concentrations, and thus independent of the applied voltage. When the voltage is changed, therefore, the accompanying concentrations changes will also be in the same proportion. For the emitter junction we thus have the relationship:

$$\frac{\Delta g(1)}{\Delta g(2)} = \frac{G_e}{G_b}. \quad \dots \dots (10)$$

From fig. 6*b*, assuming that the hole concentration curve is linear in the base, it can be deduced that:

$$\gamma = \frac{\Delta I_e^+}{\Delta I_e^+ + \Delta I_e^-} = \frac{\dfrac{\Delta g(2)}{w} D_p}{\dfrac{\Delta g(2)}{w} D_p + \dfrac{\Delta g(1)}{L_e} D_n},$$

from which, with the help of (10), we find for the emitter efficiency:

$$\gamma = \frac{1}{1 + \dfrac{G_e D_n w}{G_b D_p L_e}}. \quad . \ . \ . \ . \ (11)$$

Some requirements for emitter, base and collector

Equation (11) shows various possibilities for bringing γ close to unity. D_p and D_n are constants of the pure material of the transistor, in the present instance, germanium. L_e is made as large as possible by allowing the crystal to grow as regularly as possible; irregularities promote recombination and generation and thus reduce the diffusion length. The most important possibilities for increasing the value of γ lie in making both G_e/G_b and w small. Now a small minority equilibrium concentration necessarily implies a large majority equilibrium concentration and vice versa, and since the conductivity is determined by the majority concentration, the conductivity of the emitter should be large and that of the base small, to get a low value of G_e/G_b. The requirement that w be small (thin base) is not only to obtain a large emitter efficiency, but is, of course, also favourable for the base efficiency, since in a thin base only few holes are lost. For the same reason it is advantageous if the diffusion length in the base is long. In the base, too, the crystal must therefore be as regular as possible. The loss of the transistor action if the thickness of the base is several times the diffusion length has already been pointed out (fig. 6*a*). In such a case, the base efficiency $\beta = 0$.

The electron current across the collector is a leakage current, which is desirable to keep small; this can be done by making G_c small. Therefore the collector should also be of high conductivity.

Approximate calculation of β and α

If the loss through recombination in the base is small, and β therefore little less than 1, it is easy to deduce an expression for β. The concentration curve then deviates only slightly from the straight line in fig. 5. The number of holes in the base, summed over the length w (per cm^2 cross-section), is thus given fairly well by the area of the triangle under this concentration line and is equal to $\frac{1}{2}wg(2)$. Now minority charge carriers have an average lifetime of τ (in the base τ_b), which is proportional to their probability of recombining with a majority charge carrier. The lifetime does depend on the majority concentration though not in a simple way, owing to the intermediate steps involved in the recombination process. Apart from this, however, since we are considering concentration changes which are small compared to the equilibrium concentration of the majority carriers, and the other factors which influence the life do not change, we can consider the majority concentration as being constant and therefore the same applies for τ. The concept "average lifetime" implies that on the average, a number of holes, equal to the number present at any given moment disappear by recombination within a time τ. Thus, in the base, $\frac{1}{2}wg(2)/\tau_b$ holes disappear per second. For a variation $\Delta g(2)$, brought about by a change of voltage across the emitter junction, the loss of holes caused by recombination changes by $\frac{1}{2}w\Delta g(2)/\tau_b$. Since the number of generations per second is constant (see I), $\frac{1}{2}w\Delta g(2)/\tau_b$ is also the change in the difference between the number of holes brought in from the emitter per second and that led off to the collector. This means that:

$$\Delta I_c^+ = -\left\{\Delta I_e^+ - \frac{q}{2\tau_b}\, w\Delta g(2)\right\}.$$

Assuming again that the hole concentration in the base falls away linearly, we have again:

$$\Delta I_e^+ = qD_p\,\frac{\Delta g(2)}{w}.$$

Hence we obtain that:

$$\beta = -\frac{\Delta I_c^+}{\Delta I_e^+} = 1 - \frac{w^2}{2\tau_b D_p}. \quad . \ . \ . \ (12)$$

This formula shows the quantitative effect of the base thickness w and the lifetime τ_b on the base efficiency.

Since equation (11) is still approximately valid when there are not too many recombinations in the base, one can find from (9) an approximate formula for the current amplification factor α, viz.:

$$\alpha = \frac{1 - \dfrac{w^2}{2\tau_b D_p}}{1 + \dfrac{G_e D_n w}{G_b D_p L_e}}. \quad . \ . \ . \ . \ (13)$$

In this equation the applied voltages V_e and V_c do not appear. If we neglect the complication which arises because w is somewhat dependent on the applied voltage (Early effect), we may then consider α as a constant of the transistor. We shall see that with this assumption we can easily predict the characteristics of the transistor.

It is clear that the diffusion constant and the lifetime of the minority charge carriers are related to the diffusion length. In fact,

$$L = \sqrt{\tau D} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

This equation is frequently used in the literature on transistors and semi-conductors in general. The constant $\sqrt{\tau D}$, which plays an important part in the theory, is called the diffusion *length* because of its significance in the one-dimensional case [9]).

Substituting from (14), we can write (12) as

$$\beta = 1 - \frac{w^2}{2L_b^2}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

from which it is seen that the ratio of base thickness to diffusion length determines the efficiency of the base.

Exact calculation of α, β and γ for the one-dimensional case

To calculate α, β and γ exactly, one must start from the two equations which govern the behaviour of holes in the base. The first equation expresses the fact that the increase per second of the hole concentration at an arbitrary point is equal to the difference in the number of holes which flow in and out of unit volume per second, less the difference between the number of recombinations and generations per second, i.e.

$$\frac{\partial g}{\partial t} = -\frac{1}{q}\frac{\partial I^+}{\partial x} - \frac{g-G}{\tau} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (16)$$

The second equation expresses the fact that the hole current is caused by diffusion:

$$I^+ = -q D \frac{\partial g}{\partial x}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (17)$$

We assume here that all concentration changes take place so slowly that at any moment we are dealing with a stationary state (i.e. the analysis does not apply to high frequency changes). Then $\partial g/\partial t = 0$. Combining this with (16) and (17) gives

$$D\frac{d^2g}{dx^2} - \frac{g-G}{\tau} = 0. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (18)$$

Apart from this differential equation, g must also satisfy the boundary conditions determined by the emitter voltage and collector voltage (see fig. 6*b* and eq. 1) in the limiting planes of the base [10]).

Solution of (18) gives the hole concentration as a function of x and by differentiating with respect to x, also the concentration gradient; substituting in (17) gives the hole current I^+ as a function of x. Next, by substituting the values of x for the boundary planes *2* and *3* (fig. 6*b*), we find I_e^+ and $-I_c^+$ respectively, both as a function of V_e and V_c. Differentiating with respect to V_e at constant V_c gives ΔI_e^+ and $-\Delta I_c^+$ after which β can be calculated from (7). The result is:

$$\beta = \operatorname{sech} \frac{w}{L_b}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (19)$$

To determine γ, we must calculate $\Delta I_e = \Delta I_e^+ + \Delta I_e^-$ (see (6)). I_e^- is found directly by application of (5). Differentiation of the result with respect to V_e produces ΔI_e^-. We then find for γ:

$$\gamma = \frac{1}{1 + \dfrac{G_e D_n L_b}{G_b D_p L_e} \tanh \dfrac{w}{L_b}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (20)$$

The product of (19) and (20) gives α. For small values of w/L_b, the equations (19) and (20) reduce to the approximate equations (12) and (11) found earlier.

In practice, the situation is less favourable than would be indicated by (20). As we have already said, the diffusion length is strongly dependent on the lattice defects and certain impurities in the crystal, since the lifetime of the minority charge carriers is generally shortened by such disturbances. At the outside surface, the periodicity of the lattice is completely disrupted where the crystal ends. The recombination which occurs here sucks holes to the crystal surface. This suction is further influenced by the presence of foreign atoms adsorbed on the surface, and reinforced because mechanical processes, (grinding, scouring and polishing) have often deformed the surface. Usually, the recombination at the surface is more important than that inside the base. A special etching process is used to reduce these deleterious surface effects as much as possible.

The base current

We have seen how some of the holes injected into the base by the emitter disappear as a result of the recombination excess, while the remainder flows into the collector. There is no flow of holes to or from the base contact since, as pointed out earlier, the concentration gradient is directed substantially along the x-axis and the hole current follows this gradient. The current of holes could therefore be dealt with uni-dimensionally. This is not the case with the electron current. An electron current flows into the base from the collector and electrons are led off into the emitter. These currents are determined by the concentration gradients at the junctions in the collector and emitter respectively (where the electrons are minority carriers), and hence by the voltages across the collector and emitter junctions. In general, more electrons are led off across the emitter junction than flow into the base from the collector. Furthermore, electrons are needed to compensate for the recombination surplus. A continuous supply of electrons via the base contact is therefore needed. The current which results is called the "base current", I_b. Since the collector electron current is a (constant) saturation current, the change in the required supply of electrons which accompanies a change in emitter current must come entirely from the base current. The role which the base current plays is discussed in the next section.

---

9) The fact that, in the one-dimensional case, $\sqrt{\tau D}$ is the distance in which the deviation from the equilibrium concentration changes by a factor e, is derived in I, p. 221. A constant $1/b$ is used there in place of τ. The connection between b and the lifetime was not mentioned.

10) Equation (18) was also used for the calculation in I (p. 238) of the minority current across a *P-N* junction. The second boundary condition there was different: the concentration had to approach equilibrium values at $x = \infty$.

Here too, we have drawn conclusions about the majority charge carriers from their known behaviour in adjoining regions where they are minority carriers.

Circuit with common base and circuit with common emitter; the current amplification factor α'.

In *fig. 7a* the circuit shown in fig. 4 is reproduced, this time with the usual symbol for a *P-N-P* transistor. The emitter current I_e flows through the A.C. input source; I_e is thus the input current. The A.C. input source can also be included in the circuit in the manner shown in fig. 7*b*, the base current I_b is then the input current.

The D.C. sources used to bias the transistor, form a short circuit to the alternating currents which occur in the network. Where the A.C. characteristics of the network are concerned, then, we can leave these D.C. sources out of consideration. Thus, in fig. 7*a*, the A.C. input source e_i and the load resistance R_l are both connected directly to the base lead

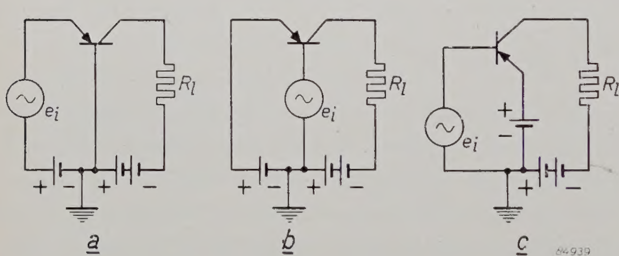


Fig. 7. *a*) Transistor circuit with common base. If the internal resistances of the D.C. sources are neglected, then A.C. input source e_i and load resistance R_l are both connected directly to the base lead of the transistor.
The *P-N-P* transistor is here indicated by the usual symbol. The emitter is distinguished from the collector by an arrow; the arrow is directed towards the crystal because the D.C. flowing to the emitter is in this direction.
b) Transistor circuit with common emitter. Disregarding the D.C. sources, the A.C. input source and the load resistance are both directly connected to the emitter lead.
c) Usual way of drawing a common emitter circuit. The circuit is identical with that in (*b*).

of the transistor. This method of connection is therefore known as "circuit with common base", or, since the common base is usually grounded, "grounded base circuit". In fig. 7*b*, neglecting again the D.C. sources, the A.C. input source and the load are both connected directly to the emitter lead of the transistor; this circuit is therefore known as "circuit with common emitter" or "grounded emitter". In this case the circuit is usually drawn as shown in fig. 7*c*.

For a circuit with common emitter, the current amplification is the ratio of the change in collector current to that of the base current, and for this circuit the current amplification factor α' is defined analogously to that for the circuit with common base, viz. [11]):

$$\alpha' = \left(\frac{\partial I_c}{\partial I_b}\right)_{V_c} \quad \ldots \ldots \quad (21)$$

Taking into account the sign convention for current (see fig. 4), we have:

$$\Delta I_e + \Delta I_c + \Delta I_b = 0. \quad \ldots \quad (22)$$

At constant collector voltage (see eq. 8), $\Delta I_e = -\Delta I_c/\alpha$; hence from (22),

$$\left(\frac{\Delta I_c}{\Delta I_b}\right)_{V_c} = \frac{\alpha}{1-\alpha},$$

so that

$$\alpha' = \frac{\alpha}{1-\alpha}. \quad \ldots \ldots \quad (23)$$

For a good transistor α is not much less than 1, and from (23) it is seen that α' is therefore much greater than 1. A reasonable value of α is for example $\alpha = 0.98$; then $\alpha' = 49$.

Input and output impedances of transistor circuits with common base and common emitter

Practical importance of input and output impedance of an amplifier

For a given amplifier and a given signal source, the aim in principle, is to obtain the maximum transference of power from the source to the amplifier and from the amplifier to the load. To achieve this, source, amplifier and load must be matched. Matching is attained when the ratio of the A.C. input voltage to the input current of the amplifier (input impedance) is equal to the internal resistance of the source of the signal, and, at the same time, the ratio of the A.C. output voltage to the output current (output impedance) is equal to the load resistance. (For simplicity we restrict ourselves to the case where all the impedances involved can be taken to be pure resistances).

Whether such matching can be achieved in practice depends largely on the order of magnitude of the input and output impedances. In view of their importance in practice, we shall qualitatively compare the impedances of transistor circuits with common base and with common emitter.

[11]) The quantity α' thus defined is always positive: the addition of a minus sign in the defining equation is therefore unnecessary, in contrast with the case of the definition of α (see eq. 8).

Input and output impedances of transistors

For a transistor, in fact for amplifiers in general, it is not possible to quote an input and output impedance without further qualification. The input impedance depends on the load and the output impedance on the internal resistance of the signal source connected to the input side.

The output impedance in a circuit with common emitter is lower than with common base. To appreciate this, we must take into account that the concentrations of the minority charge carriers at the collector are not exactly zero, but that these concentrations depend somewhat on the voltage across the collector junction. In order to determine the output impedance, it is necessary to be able to vary the voltage between collector and base; an A.C. source is therefore connected between them (*fig. 8*). It is

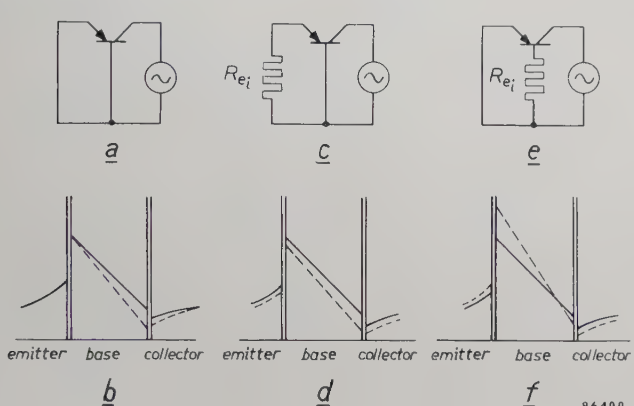


Fig. 8. As a result of the internal impedance of the signal source (here, for simplicity, represented as a pure resistance R_{ei}) the output impedance of a transistor with grounded base (*c* and *d* is higher, with grounded emitter (*e* and *f*) lower than when base and emitter are short-circuited for A.C. (*a* and *b*). The concentration distributions of the minority charge carriers corresponding to each circuit are shown at a moment when the voltage across the collector junction is slightly more in the reverse direction (broken lines) than that due to the simple D.C. bias (full lines). To simplify the diagrams the D.C. sources providing the bias are not drawn.

useful to begin the argument from the hypothetical situation represented in fig. 8*a*: here the signal source is assumed to be of zero internal resistance so that no voltage variation due to the generator in the collector circuit can appear between emitter and base. Hence the minority concentrations at the emitter remain fixed. Let us suppose that the voltage across the collector changes in such a way that the minority concentrations at the collector drop a little: an increase then occurs in the current across the collector (see fig. 8*b*, in which, for simplicity, recombination and generation in the base have been neglected). The hole contribution to this increase (the hole concentration gradient has increased slightly) necessitates an extra current via the emitter circuit; the electron contribution requires an extra current via the base circuit. Both extra currents are so directed that positive charge flows towards the crystal. Since the concentration changes at the collector are small, the changes in current are small, i.e. the output impedance is high.

In practice, in the circuit with *common base*, the internal resistance R_{ei} of the signal source appears in the emitter circuit (fig. 8*c*). The extra current in the emitter circuit causes a voltage drop in this circuit which reduces the potential difference between emitter and base. The concentrations at the emitter thus drop and the increase in the hole concentration gradient in the base is partly compensated (fig. 8*d*). The hole current thus alters less than was the case in fig. 8*b*. The change in the electron current over the collector junction has remained the same, so that the total collector current change has decreased, i.e. the output impedance is even higher than with the hypothetical short-circuited emitter circuit.

For a circuit with *common emitter*, the internal resistance of the signal source is in the base circuit (fig. 8*e*). The extra current in the base circuit (due to the voltage change across the collector junction) results in a voltage drop which serves to increase the potential difference between emitter and base, and this causes the concentrations at the emitter to rise. The original increase in the hole current as a result of the fall in concentration at the collector, is thus reinforced (fig. 8*f*). The collector current therefore also increases and the output impedance is lower than with short-circuited base circuit.

The emitter electron current also increases. Hence the whole increase in the collector electron current does not have to be carried away via the base circuit, for some of it disappears across the emitter junction. The voltage across the emitter junction changes until the balance between the electron transport to and from the base is restored.

The greater the value of R_{ei}, i.e.the more difficult it is to lead electrons away via R_{ei}, the greater is the portion of the electron current which finds its way across the emitter. The voltage difference across the emitter junction must therefore adjust itself to a correspondingly higher value; hence the rise in the hole concentration at the emitter, and thus the hole current through the base, increase still further. This means that as R_{ei} is increased, the output impedance decreases.

We have here a very remarkable situation. The current generated in the collector circuit by a source, divides itself between two parallel branches, viz. the emitter and the base. If R_{ei} is placed in the

emitter branch, the result of an increase in R_{ei} is an increase in the effective resistance of the parallel combination, as would be expected. However, if R_{ei} is in the base branch, the effective resistance becomes *less* as R_{ei} becomes greater, which is quite contrary to the normal behaviour of parallel resistances.

The output impedance with common base can easily be several megohms in practice; with common emitter some ten thousands of ohms is a normal value.

The *input impedance* of a transistor in a common base circuit is considerably lower than that in a common emitter circuit (e.g. 25 to 30 Ω for the former and 1200 Ω in the latter case). For a given alternating input voltage we have to deal in the first case with fluctuations of the emitter current, which are much larger than those of the base current; the latter, however, determine the input impedance in the second case.

Comparison of transistors with thermionic valves

Transistors and amplifying valves exhibit considerable differences in their properties even at audio frequencies. This is largely due to the fact that the input and output impedances of transistors differ so much from the values usually occurring for valves. In a pentode, for example the input and output impedances are generally very large with respect to the internal resistances that are usual for signal sources and load resistances used in practice. In most cases, therefore, the A.C. output i_a of a pentode is given by $i_a = Sv_g$, where S represents the mutual conductance and v_g the e.m.f. of the signal source. The mutual conductance thus determines the behaviour of the pentode; it makes little difference how large the input and output impedances are, provided they be large.

For a transistor too, the mutual conductance can be found. With common base it is

$$S = \left(\frac{\partial I_c}{\partial V_e}\right)_{V_c};$$

With common emitter, the same expression applies, but with a minus sign, since an increase of the input voltage in the latter case means a decrease of the voltage V_e across the emitter junction.

Now:

$$\left(\frac{\partial I_c}{\partial V_e}\right)_{V_c} = \left(\frac{\partial I_c}{\partial I_e}\right)_{V_c}\left(\frac{\partial I_e}{\partial V_e}\right)_{V_c} = -\alpha\left(\frac{\partial I_e}{\partial V_e}\right)_{V_c}.$$

If one assumes that the hole concentration in the base is linear, it is possible with the help of fig. 6*b* and equation (1) to deduce an expression for I_e as function of V_e; by differentiating this one finds that for normal biassing of transistors ($g(l) \gg G_e$):

$$\left(\frac{\partial I_e}{\partial V_e}\right)_{V_c} = \frac{q}{kT} I_e.$$

Ignoring the sign, the mutual conductance of a transistor is thus:

$$S = \alpha \frac{q}{kT} I_e.$$

Apart from the direct current (I_e) through the emitter junction, S depends only on α. For a good transistor α is almost equal to 1, and all transistors have thus the same mutual conductance for the same D.C. biassing. For a normal value of I_e, for example 1 mA, the mutual conductance is 40 mA/V, which is much greater than can be achieved with pentodes.

However, a comparison of a transistor with a tube simply on the basis of the mutual conductances is of little value, because for a transistor the behaviour is largely governed by the input and output impedances. In a circuit with common base, the output impedance is very high (as for a pentode), but the input impedance is then only a few tens of ohms. The latter, for constant V_c, is

$$\left(\frac{\partial V_e}{\partial I_e}\right)_{V_c},$$

and this, as we have seen above, is equal to kT/qI_e; at room temperature and for $I_e = 1$ mA, its value is 25 Ω. In the circuit with common emitter the input impedance is higher, but the output impedance on the other hand is much lower than that of the circuit with common base. However, a transistor in the circuit with common emitter behaves much more like a valve than it does in the circuit with common base. For a very large value of α' (very small base current) the input inpedance would be comparable with that of a valve, but the output impedance in this case would be very low.

The behaviour of transistors is clearly more complicated than that of valves. Account must be taken of the input and output impedances, and this is complicated by the fact that these impedances are not dependent solely on the transistor: as pointed out above, the input impedance depends on the load and the output impedance on the internal resistance of the signal source.

An obvious advantage of transistors as compared with valves is that for amplification of small signals, a small bias is sufficient. For small signals, the bias with valves (i.e. anode voltage) is always ten or more times greater than the collector-base voltage of the transistor, so that the energy dissipation is much greater; transistor action is thus more efficient for small signals. Furthermore, a transistor needs no filament or heater current, an advantage which speaks for itself.

Transistor characteristics according to the simple theory

A transistor is generally characterized by a set of graphs which give the collector current I_c as function of the collector voltage V_c, for various values of the input current. The latter is the emitter current I_e with a common base and the base current I_b with a common emitter circuit. We are of course interested only in that section of the graphs which corresponds to a collector voltage in the reverse direction and an emitter voltage in the forward direction.

Common base

If the emitter current is zero (emitter open-circuited) a reverse current I_{cl}) (l for "leakage") flows through the collector, which for increasing reverse potential approaches exponentially to a saturation value I_{co}) (*fig. 9a*). From the definition of α (see eq. 8) it follows, by integrâtion:

$$I_c = I_{cl} - \alpha I_e. \quad . \quad . \quad . \quad . \quad (24)$$

For a *P-N-P* transistor, which we have been

considering throughout, I_e is positive, but I_c and I_{cl} are negative, like V_c (reverse potential). In the graph we therefore plot $-I_c$ as a function of $-V_c$ for the various values of I_e.

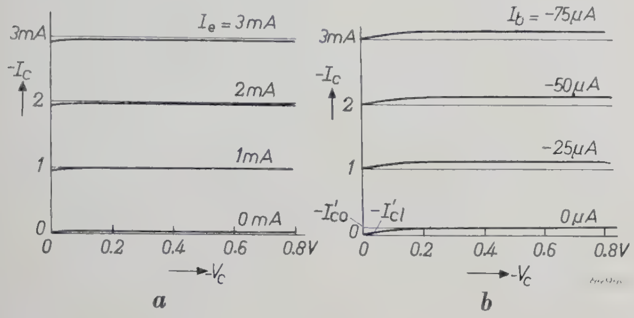


Fig. 9. Theoretical characteristics of a transistor: collector current I_c as function of the collector voltage V_c for various values of the input current. According to the sign convention (fig. 4) V_c and I_c are negative: $-V_c$ and $-I_c$ have therefore been plotted.
a) For a transistor with common base circuit. The lowest characteristic gives I_{cl}, the collector current with open-circuited emitter.
b) For a transistor in circuit with common emitter. The parameter here is the base current I_b. It should be noted that all the values of I_b shown are negative. I'_{cl} is the saturation value of the collector current I'_{cl} with open-circuited base.

The characteristic for any value of I_e is now obtained by moving the characteristic for $I_e = 0$ upwards by an amount αI_e.

Common emitter

If the base current is zero (base open-circuited), a current I'_{cl} flows which, for the same collector voltage is a factor $(1 + \alpha')$ larger than the leakage current I_{cl} in the case of the common base. This may be seen as follows. We have

$$I_e + I_b + I_c = 0.$$

Adding this to (24), we can write

$$(1-\alpha)\, I_e = -I_b - I_{cl}\,.$$

For open base circuit, $I_b = 0$ and $I_e = -I'_{cl}$ so that

$$(1-\alpha)\, I'_{cl} = I_{cl}\,.$$

From (23) it follows that $1 - \alpha = 1/(1 + \alpha')$, so that

$$I'_{cl} = (1 + \alpha')\, I_{cl} \quad . \quad . \quad . \quad . \quad . \quad (25)$$

I'_{cl} also approaches exponentially to a saturation value I'_{co} at increasing reverse voltage and, following directly from (25) we have

$$I'_{co} = (1 + \alpha')\, I_{co}\,. \quad . \quad . \quad . \quad . \quad (26)$$

From (21) we obtain by integration:

$$I_c = I'_{cl} + \alpha' I_b. \quad . \quad . \quad . \quad . \quad . \quad (27)$$

For a *P-N-P* transistor I_c, I'_{cl} and I_b are all negative. We therefore plot $-I_c$ as a function of $-V_c$ for the various values of I_b.

The characteristic for any value of $-I_b$ is obtained by shifting the characteristic for $I_b = 0$ upwards by an amount $\alpha'(-I_b)$ (fig. 9*b*).

If the characteristics for common base and common emitter are plotted on the same scale, then, as follows from (25), at the same value of $-V_c$ the characteristics will have a gradient for common emitter which is $(1 + \alpha')$ times that for common base. The gradient of the characteristics represents the internal reciprocal resistance of the transistor at constant input current, i.e. using a signal source of infinite internal resistance. In this case the output impedance for common base is thus $1 + \alpha'$ times as great as that for common emitter.

Early effect

The actual shape of the transistor characteristics differs somewhat from the shape deduced above, as a result of the *Early effect* which was mentioned on p. 240. Up till now we have assumed that the boundary planes of a *P-N* junction, e.g. the sections *1* and *2*, or *3* and *4* in fig. 6*b*, had a fixed position in the crystal. In reality, the space charge region which exists between the two planes (see p. 234) becomes thicker as the voltage in the reverse direction increases. This results principally from a shifting of the plane which lies on that side of the *P-N* junction where the specific resistance is greatest. In a transistor it is plane *3* which undergoes the largest displacement because this plane is situated in the base which, as we have seen, usually has a greater specific resistance than the collector. With increasing reverse voltage across the collector, section *3* shifts towards the emitter and the effective base thickness decreases. This is known as the Early effect [12]). This makes α, and hence also α', larger. The transistor characteristics will therefore be less flat, and the output impedances at constant input current will be smaller, than would be expected from the simple theory. At section 2 on the emitter side of the base, the same effect occurs. There it is much less important, since the voltage variations across the emitter junction are so much smaller than those across the collector junction.

A further important consequence of the Early effect is that with increasing collector voltage, i.e. with increasing α, the emitter current increases. There is thus a feedback effect of the output voltage on the input current, which is undesirable. To limit this effect, it is necessary that the specific resistance of the material of the base is not too large.

P-N-P and *N-P-N* transistors

In unfolding the theory, we have always referred to *P-N-P* transistors. For *N-P-N* transistors — where the base is of *P*-germanium and emitter and

[12]) J. M. Early, Effects of space-charge layer widening in junction transistors, Proc. I. R. E. **40**, 1401-1406, 1953.

collector of N-germanium — the theory is completely analogous. In principle, the theory can be developed without mention of holes and electrons, by dealing simply with minority and majority charge carriers: the theory then becomes identical for both types of transistor. Practical differences in the behaviour of the two types are caused by the fact that the diffusion constant for electrons is about twice that for holes.

Summary. The action of a junction transistor is based on the properties of P-N junctions in a germanium crystal. Such a transistor is built up of a P-N junction from "emitter" to "base" followed at a very small distance by an N-P junction from "base" to "collector". The action is explained, starting from the fact that the current at a junction is determined by the diffusion currents of the minority charge carriers which can easily be calculated at the boundary planes of such a junction. It is pointed out that the voltage applied across a junction determines the concentrations of the minority charge carriers in the boundary planes. At one of the junctions this concentration is fixed at zero; a small change of voltage across the other junction then results in a strong variation of the concentration gradient, and thus of the diffusion current of the minority charge carriers, between the two junctions.

Emitter and base efficiency are discussed and calculated for the one-dimensional case; this is also done for the current amplification factors for circuits with common base and common emitter. The difference in the input and output impedances for these two circuits is explained and the characteristics which follow from the theory, are deduced.

A MACHINE FOR BEND TESTS

620.177.3

A requirement frequently imposed upon a metal is that a specimen of certain form should not crack or break when bent through a prescribed angle around a mandrel of prescribed diameter. The result of a bend test, however, depends upon the manner in which the test is carried out, and this is not always specified. The consequence may be a difference of opinion between supplier and customer as to whether the metal satisfies or falls short of the specification.

With a view to overcoming such difficulties the Materials Research Group of the Research Laboratory in Eindhoven has developed a special bending machine for testing sheet and strip metal (*fig. 1*). By means of a single hand-operation the test specimen can be bent in this machine quite smoothly and with the correct radius through any angle between 0 and 180°. The design is shown in *fig. 2*. The test specimen (*4*) is held between the clamping piece (*1*) and the mandrel in the form of an interchangeable radius plate (*3*), which is accurately rounded on the top edge. The upper face of the clamping piece is shaped in such a way as to support the specimen as close as possible to the bending point, thus preventing unwanted deformation. The bend is effected by the thrust block (*7*), which is moved by the lever (*9*). Three rollers (*6*) mounted in the thrust block ensure that, when the lever is pushed over, the block slides along the specimen almost without friction so as to exert practically no tensile load upon it. This is important, for if there were any appreciable tensile load on the specimen it would not be subjected to a pure bend and premature cracking might occur. Allowance can be made for the thickness of the specimen by adjusting with wing-nut *8* the position of the thrust block in its groove on the lever. The lever *9* turns about an axis which coincides with the centre of curvature of the radius plate; this is essential to the correct functioning of the machine. The interchangeable radius

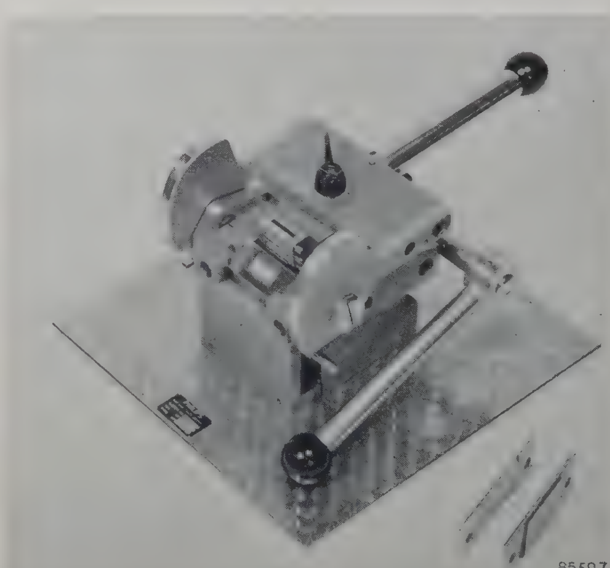


Fig. 1. The newly developed machine for testing sheet or strip metal for its bending properties. The test specimen is inserted between rollers and a radius plate. At the right can be seen two of the interchangable radius plates over which the specimen is bent.

plates are ground in such a way that the centre of curvature always lies at the same position in the machine. A support bar (*2*) takes up the horizontal component of the force exerted upon the radius plate during operation and thus prevents it from bending under the strain. When the angle of bend exceeds 90° this component reverses direction and the clamping piece (*1*) then takes up the horizontal

force. The support bar is then no longer necessary and, indeed, it forms an obstruction to further bending. For this reason it is pressed down against two springs (*12*) by the cams (*5*) on the end of the lever (*9*), which bear on both ends of the pin (*11*) when the angle of bend becomes greater than 90° (fig. 2*b*). In this way the bend test can be continued unhindered up to 180°. The angle of bend can be read from one of the cams (5) which is calibrated in degrees. Even at an angle as small as 60° to 70°, the outside of the bend becomes visible with the specimen still in position in the machine. It can thus readily be perceived at what angle of bend crack formation begins.

Fig. 3a shows the cross-section of a specimen of chrome steel with a tensile strength of approximately 65 kg/mm^2 and a thickness of 1.9 mm, which has been bent through 180° over a diameter of 0.95 mm. The diameter of the radius plate in this case was therefore half the thickness of the specimen. It can be seen that, even under such very adverse conditions for a bend test as these, the specimen nevertheless follows the radius plate quite faithfully. Fig. 3*b* shows that under more favourable conditions (mild steel with a tensile strength of 40 kg/mm^2 and a thickness of 1.0 mm, bent over a diameter of 1.20 mm) the inner profile follows almost exactly the curvature of the radius plate.

As stated, the machine described serves for testing sheet or strip metal, but it is also possible to design a machine on the same principles for testing metal in wire or bar form.

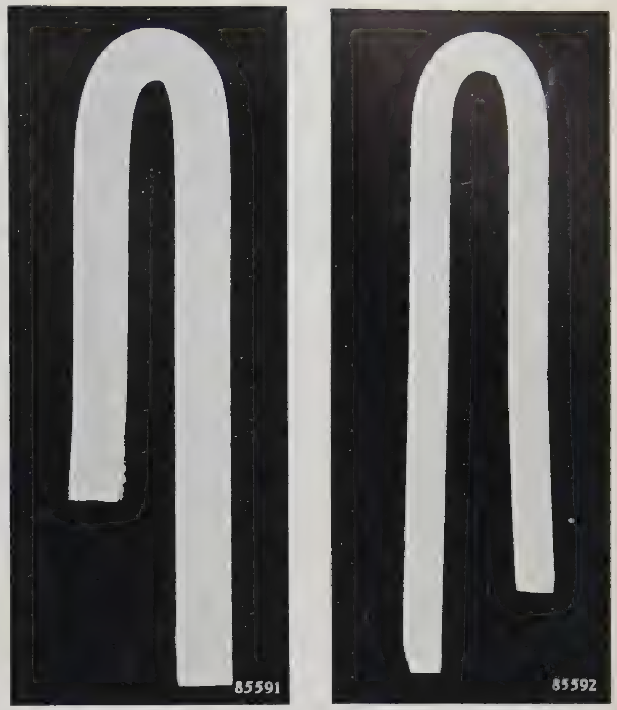


Fig. 3. *a*) Section of a test specimen of chrome steel, bent through 180° by the machine shown in fig. 1 (tensile strength of specimen approx. 65 kg/mm^2; specimen thickness d = 1.9 mm; diameter of radius plate D = 0.95 mm; thus D/d = 0.5).
b) Section of a test specimen of mild steel (tensile strength approx. 40 kg/mm^2; d = 1.0 mm; D = 1.2 mm, and D/d = 1.2).

In order to make clear why this rather elaborate machine has been developed for performing such a simple test, some comments are called for on the rather primitive methods often employed for the bend test.

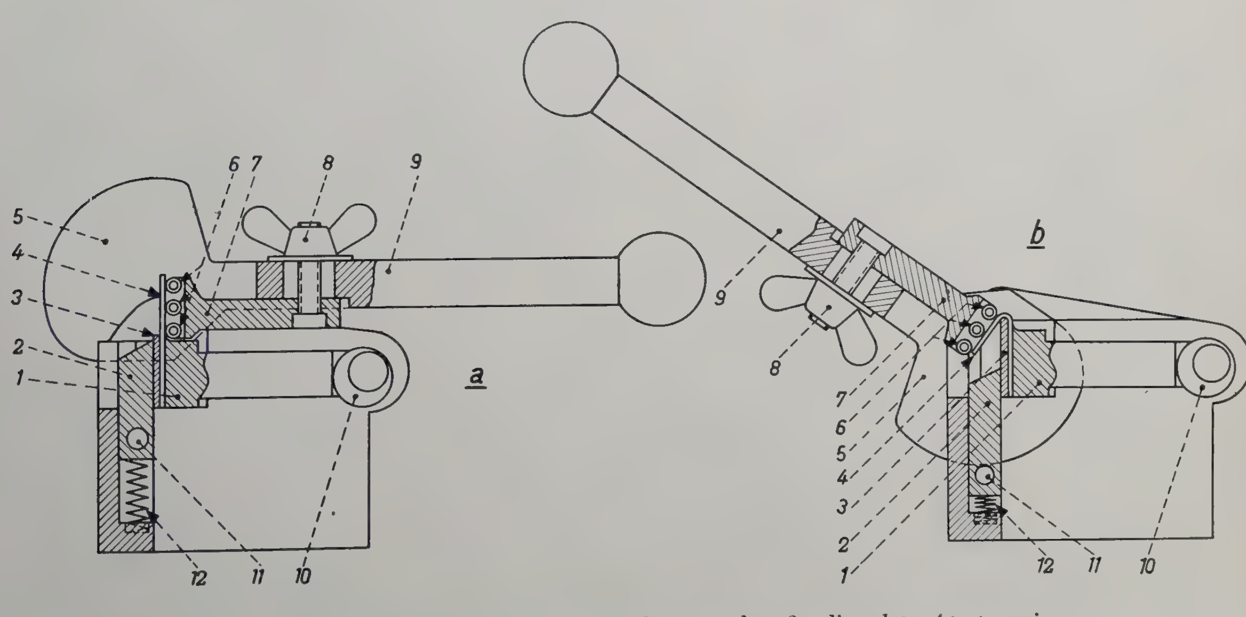


Fig. 2. Structural design: *1* clamping piece; *2* support bar; *3* radius plate; *4* test specimen; *5* cam, which depresses support bar; *6* rollers, mounted in thrust block *7*, which is fixed by wing-nut *8* to lever *9*; the excentric *10* serves to hold clamping piece against specimen; *11* pin by which the support bar is depressed against spring *12*. *a*) Machine ready for operation after insertion of test specimen; *b*) after bending through 145°.

A method in common use is to secure the test specimen together with a mandrel in a vice (*fig.4a*) and to bend the specimen over the mandrel by hand or with a hammer. If the specimen is to follow the mandrel closely, the bending force must be applied close to the mandrel (fig. 4*b* and *c*). The use of a hammer is to be deprecated, as it subjects the specimen to shock-loading of unknown magnitude, which may result in premature cracking or breaking. What is more, hammering close to the mandrel may produce undesirable work hardening of the regions subjected to impact.

If an angle of bend greater than 90° is specified, it is necessary, when using this simple method, to remove the test specimen at a 90° bend from the

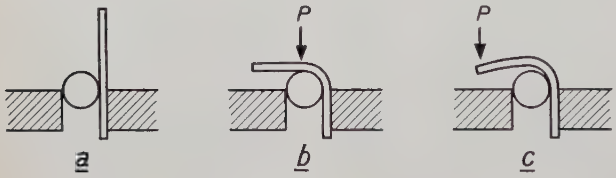


Fig. 4. Specimen and mandrel clamped together in a vice for a primitive bend test (*a*). If the bending force *P* is applied near to the mandrel, the bend radius follows the curvature of the mandrel (*b*); if *P* is applied at some distance from the mandrel, the bent radius will be too large (*c*).

vice, to secure it again in a different way (*fig. 5a*) and then to continue the bend test by tightening the vice itself. It is now almost impossible to bend the specimen closely around the mandrel, because the bending forces are necessarily applied at some distance from the mandrel, and the mandrel always slides back a little as the angle of bend increases. The consequence is a sharper bend than specified (fig. 5*b*). The second objection might be overcome by supporting the specimen directly beneath the mandrel (fig. 5*c*), but then the metal is no longer subjected to a pure bending load. Moreover, the force exerted by the support may be so considerable as to cause unwanted deformation of the specimen at the bend radius. Nor is it possible with this method to determine the point at which cracks begin to form.

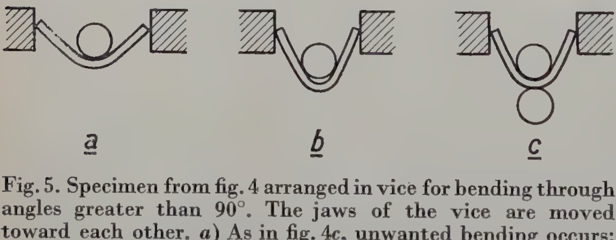


Fig. 5. Specimen from fig. 4 arranged in vice for bending through angles greater than 90°. The jaws of the vice are moved toward each other. *a*) As in fig. 4*c*, unwanted bending occurs; the mandrel also becomes displaced (*b*) causing too sharp a bend. Supporting the mandrel (*c*) avoids displacement but the specimen is then no longer subjected to a pure bending.

If an angle of bend of 180° is specified, a method frequently used is to press the test specimen into a die by means of a punch with a curvature of the prescribed radius. The aperture of the die is then equal to the thickness of the punch plus twice the thickness of the specimen (*fig. 6*). In this case an unwanted load is exerted upon the specimen by the friction occurring between it and the die. This

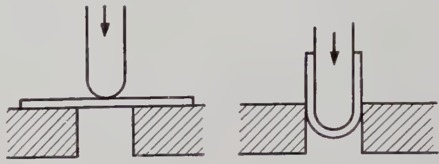


Fig. 6. Bend test employing a die and a radius plate in the form of a punch.

friction can be limited by rounding-off the edges of the die and possibly by providing the aperture with a taper, but these measures cannot be taken far without entailing the drawbacks illustrated by fig, 5*a* and *b*. The method has no practical value if the punch is thin in relation to the test specimen or if punch and specimen are both thin. In the first case the strains set up in the punch are so great as to cause it to collapse. In the second case the specimen suffers deformation as a result of the shearing forces arising.

Use is sometimes made of special tools incorporating several of the elements of the machine described here. The test specimen is secured in a clamp, one side of which is given the curvature over which the specimen is to be bent, that is to say one clamping piece acts as the mandrel (*fig. 7*). The metal is bent

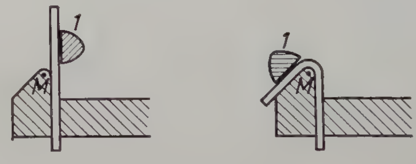


Fig. 7. Bend test employing special tools. The axis of rotation of thrust block *1* coincides with the centre of curvature *M* of the curved member over which the specimen is bent.

over this surface by means of a thrust block attached to an arm, whose axis of rotation coincides with the centre of curvature of the mandrel. With this arrangement, however, the specimen cannot be bent through more than about 110° (depending upon the requisite strength of the mandrel). Moreover a tensile load is exerted by the thrust block as it moves over the face of the specimen. These drawbacks have been substantially eliminated in the apparatus described here.

M. J. W. GEENEN.

THE "NORELCO" COUNTING-RATE COMPUTER

by E. A. HAMACHER † *) and K. LOWITZSCH *).

621.317.79:621.374.32:548.734

Analogue computers have acquired considerable importance in many fields. The computer described in this article is very simple with regard to the underlying mathematics. It may be regarded as a useful example to illustrate some of the mechanical and electrical intricacies involved, when a high accuracy is asked of this type of instrument.

Introduction

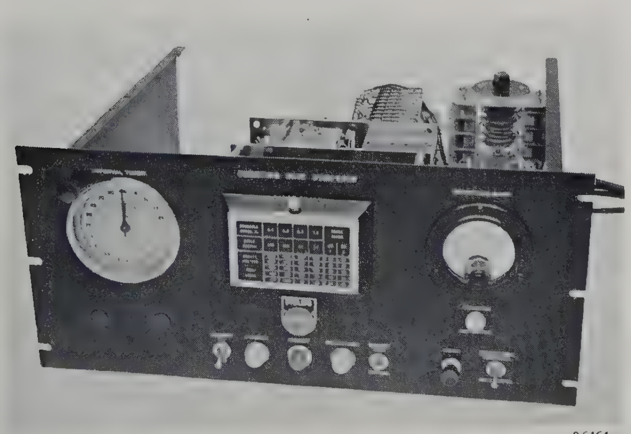
In X-ray diffractometry with the "Norelco" equipment [1]) the counting-rate computer is used for the measurement and the automatic recording of the diffracted X-ray intensity vs. Bragg angle according to the *fixed count* method. As was explained in a former article [2]) this method consists in measuring the time t necessary to accumulate a predetermined number of counts N, whereupon the counting rate n (which is proportional to the intensity of the radiation entering the detector) is computed from the formula

$$n = N/t. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The fixed count method has the advantage that each intensity value in a pattern has the same statistical relative error (expressed, for example, by the "probable" relative error ε_{50}, cf. II).

Several counting-rate recorders for automatic fixed count measurements have been developed [3]). In these instruments, however, counting rates n are recorded on an inverse scale in one case and on an approximately logarithmic scale in another. The basic idea of the "Norelco" counting-rate computer was to obtain a recording of n on a *linear* scale by automatic evaluation of eq. (1).

The computer, which is mounted in the electronic circuits and recorder rack of the "Norelco" diffractometer, is shown in *fig. 1*.



86161

Fig. 1. Chassis and front panel of the "Norelco" counting-rate computer. This unit is placed in the electronic circuits rack of the diffractometer. The computer output can either be fed to the recorder or be read on the meter at the upper right. At the upper left a register is mounted which is not used for the computer but only in manual fixed count measurements (cf. II).

Principle of the computer

The analogue to equation (1) which is used as the basis of the computer, is Ohm's law:

$$I = E/R \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

The sliding contact of a slide-wire resistor is moved by a synchronous motor, so that the resistance R of the used portion of wire will increase with time at a constant selectable rate k:

$$R = kt \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

The current I flowing in the resistor when a fixed voltage E is applied to it will gradually decrease. When the time t in eq. (3) is made equal to that in (1), the current I evidently will be proportional to the counting rate n, and therefore may be directly fed to the strip-chart recorder to give the desired linear recording.

To make use of the analogue for the diffractometer, thus amounts to the following procedure. At a given Bragg angle position of the detector, the movement of the contact on the slide-wire is automatically

*) Philips Laboratories, Irvington-on-Hudson, N.Y., U.S.A.

1) W. Parrish, E. A. Hamacher and K. Lowitzsch, The "Norelco" X-ray diffractometer, Philips tech. Rev. **16**, 123-133, 1954/1955 (No. 4). This article is hereafter referred to as I.

2) W. Parrish, X-ray intensity measurements with counter tubes, Philips tech. Rev. **17**, 206-221 1955/56 (No.7-8). This article is hereafter referred to as II.

3) S. W. Lichtman, E. T. Byram and H. Friedman, Strip chart recording with an autoscaler, Electronics **23**, April 1950, page 122. Another instrument was designed by the Berkeley Scientific Corp., Richmond, Calif.

started when counting begins, i.e. when the scaling circuits of the diffractometer are put into operation (cf. II, fig. 12). The movement is stopped and a fixed voltage is applied to the slide-wire when the predetermined number of counts has been accumulated, i.e. after one revolution of the mechanical register of the diffractometer has been completed (the accumulated number of counts N is then 100 $\times$ the selected scaling factor, thus 100, 200, . . ., 12 800 or 25 600). The current now flowing in the slide-wire is recorded on the strip chart. After a measurement has been made and recorded, the sliding contact is reset to its initial position, the detector is moved by the goniometer to the next Bragg angle position and the operation is repeated. A continuous trace consisting of steps of equal width is thus obtained; see *fig. 2*.

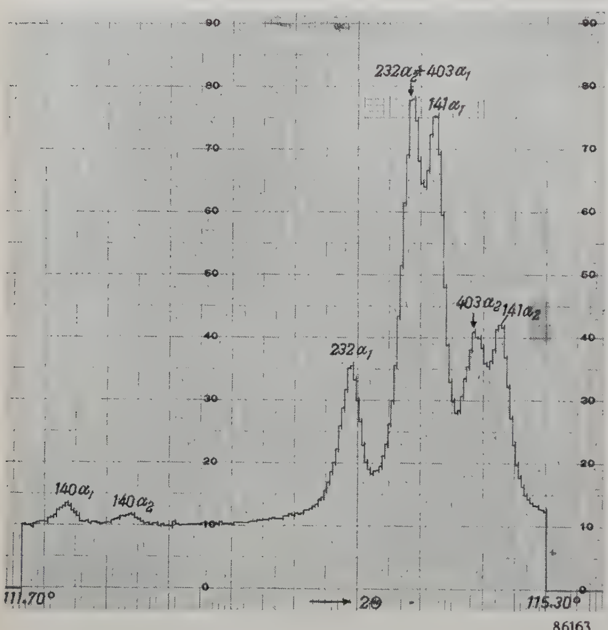


Fig. 2. Chart of diffraction pattern recorded by the "Norelco" counting-rate computer. The Bragg angle steps in this instance are 0.02°(2Θ); the statistical error ε_{50} of each point is 0.8%. This recording took about 3 hours of completely automatic operation.

In order to understand the conditions involved in using the analogue, it is important first to consider the limiting case of the measurements.

Let the speed of the sliding contact, i.e. the rate k of increase of R, be given. The distance the sliding contact has travelled when N counts have been accumulated, will be larger the smaller the intensity (counting rate n) which is being measured. In fact, from eq. (1) and (3) it follows that

$$R = k\frac{N}{n}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

The *minimum* measurable counting rate, $n_{\min}$, will be that for which the whole length of wire, $R = R_{\max}$, is used; the corresponding time will be largest:

$$n_{\min} = \frac{N}{t_{\max}} = \frac{Nk}{R_{\max}}. \quad . \quad . \quad . \quad . \quad (5)$$

On the other hand, in order to protect the recorder, the current fed to it must not exceed a certain permissible value, $I_{\max}$. The *maximum* recordable counting rate, $n_{\max}$, will be the one for which the fixed number of counts N is accumulated in such a short time ($t_{\min}$) that the used portion of the slide-wire will just have attained the minimum resistance value, $R_{\min}$, required to keep the current to $I_{\max}$. Thus:

$$n_{\max} = \frac{Nk}{R_{\min}} = \frac{Nk\,I_{\max}}{E}. \quad . \quad . \quad . \quad (6)$$

From (5) and (6) it is seen that the ratio

$$\frac{n_{\max}}{n_{\min}} = \frac{I_{\max}\,R_{\max}}{E}$$

is a design constant and does not depend on the values of k and N. In the actual instrument $n_{\max}/n_{\min} = 118$.

It should, of course, be possible to adjust the range ($n_{\max}$) of the instrument, and this is readily done by varying k, the speed of the sliding contact: with increased k, the total resistance of the slide wire, $R_{\max}$, according to eq. (5) will correspond to a higher counting rate $n_{\min}$. Owing to the constant ratio $n_{\max}/n_{\min}$, the range $n_{\max}$ will then be higher, too.

If N is increased in order to improve the accuracy of the measurements (i.e. to reduce the statistical error in counting, cf. II), $n_{\max}$ and $n_{\min}$ are increased (see eqs. 5 and 6). It will then be necessary to decrease k in the same proportion if the range ($n_{\max}$) is to be kept constant.

It was stated above that the resistance of the slide-wire must have attained a certain minimum value, $R_{\min}$, before the voltage E can be applied to it and the current flowing through it can be fed to the recorder. A way to ensure this would be to enforce a minimum time, $t_{\min}$, during which counting must be effected and the sliding contact moved on the wire. Since however no intermediate positions of the contact in this part of its travel would ever be used, it is simpler to substitute a resistor with fixed resistance $R_{\min}$ for this part of the slide-wire. The sliding contact (or rather its driving mechanism) whose movement is started together with the counting, should then travel a distance $\Delta = R_{\min}/R_1$ before it reaches the slide-wire and starts to move

on it; R_1 denotes the resistance per unit length of the wire. It will be noted that the "pre-travel" Δ is not affected by a change of k, whereas the time $t_{\min}$ necessary for traversing it clearly is: $t_{\min} = R_{\min}/k$ (cf. eq. 3).

Mechanical design of the computer

In the actual instrument, a 15 turn helical slide-wire ("Helipot") is used, along which the sliding contact is moved at a constant angular speed. *Fig. 3* shows a schematic drawing of the complete

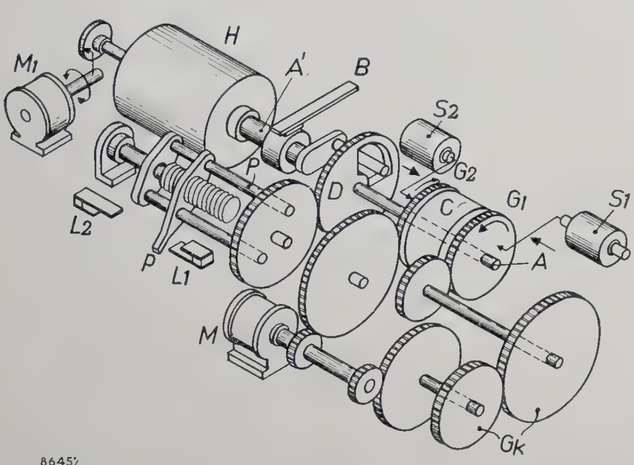


Fig. 3. Simplified drawing of the mechanical assembly. The electric motor M runs continuously, driving the gears G_1 and G_2, idling on shaft A in opposite directions. Shaft A carrying the double-face clutch drum C can be coupled either to G_1 or G_2 by the action of solenoids S_1 and S_2. H is the helical slide-wire resistor ("Helipot"), whose shaft A' is driven by A via an angular pre-travel device D; B friction brake, M_1 motor for the retractable contact mechanism (see fig. 5). P positioning device with limit switches L_1, L_2 actuated by arm p. G_k exchangeable range gears controlling the speed of movement of the sliding contact of H during the counting interval.

mechanical assembly. The 60 r.p.m. synchronous motor M turns continuously and drives the gears G_1 and G_2 idling on shaft A in opposite directions. The "Helipot" H can be coupled to either of these gears by means of the double friction clutch assembly C actuated by solenoids S_1 and S_2 respectively. Gear G_2 is used to reset the slide-wire contact to its starting position after each measurement; it turns at a constant speed of $\frac{1}{2}$ rev/sec so that the resetting will take 30 seconds if the complete length of the 15-turn "Helipot" was used ($n_{\min}$). Gear G_1 is used to drive the slide-wire contact during the counting period; it can turn at different speeds, dependent on the selected range gears G_k, providing for different rates of change (k) of the resistance. Three pairs of range-gears are available for the computer, permitting the selection of five gear ratios and, hence, five ranges; for a scaling factor of 256 ($N = 25\,600$) these correspond to $n_{\max} = 200$, 400, 800, 1600 and 3200 counts/sec.

Gear G_1 is coupled to the shaft A by the clutch at the moment when counting is started. Shaft A, however, will entrain the "Helipot" shaft A' only after it has rotated through a certain constant angle a, owing to the interposed angular pre-travel device D. The pre-travel angle a is the equivalent of the distance Δ introduced at the end of the preceding section.

The interposition of the pre-travel device has a curious consequence. In resetting the slide-wire contact, the resetting movement must be stopped by de-energizing the solenoid S_2 at the exact moment the contact has reached its initial position. It would seem natural to ensure this by arranging for a limit switch to be operated by the contact itself. Provision must be made, however, for the possibility of the counting rate being so high that in the counting operation the solenoid S_1 is de-energized and the shaft A uncoupled before the pre-travel angle has been traversed. The sliding contact has then not moved at all (the recorded intensity remaining at the top of the range) the and the limits witch could not be operated. Thus, no resetting of the shaft A would occur and in the *following* counting operation the pre-travel angle would be smaller than the required value.

For this reason — and for other reasons which cannot be discussed here — a separate positioning device P is provided which is rigidly and permanently coupled to shaft A by a pair of gears, see fig. 3. The arm p of this device describes a helical path exactly similar to that of the sliding contact of the "Helipot", except for the fact that it is longer by an angle a (distance Δ). On resetting the "Helipot", the arm p finally comes to rest on a stop assuring the accurate starting position. When touching the stop, the arm actuates a switch which disengages the reset clutch. In order to prevent the arm from recoiling from the stop, the switch turns on a time delay relay arranged to keep the reset clutch closed for about another $^1/_2$ sec: the clutch slips momentarily and the recoil energy is absorbed by the clutch and the motor M. The arm is then kept at the stop by a spring.

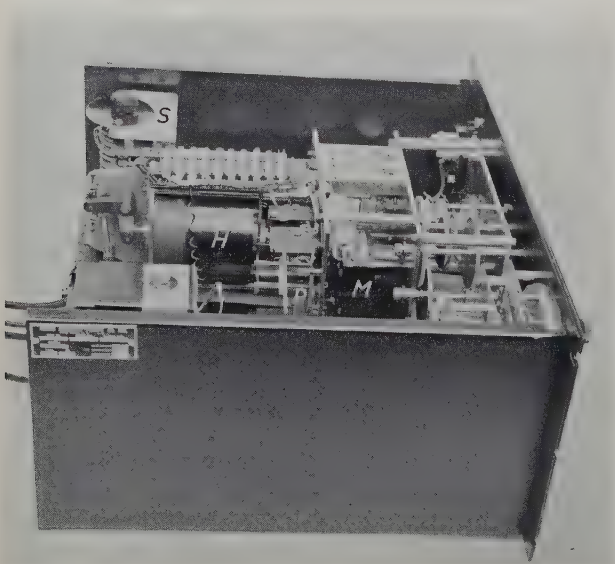
Immediately after resetting, the "Helipot" is ready for another measurement (see below). However, the synchronous motor M will have dropped back in phase a little bit (about $^1/_{15}$ of a cycle) owing to the extra load caused by the slipping clutch, and this would endanger the exact synchronism between counting time and rotating time of the "Helipot" (eq. 1 and 3). The limit switch of the positioning device therefore operates another time delay relay which prevents the starting of another counting

period for about 1 sec, by which time the motor will have caught up.

Throughout the resetting operation, the pin in the angular pre-travel device remains in contact with the driving member, and it is prevented by a friction brake B from leaving it when the resetting is completed. Thus the shaft A always has to travel through the complete pre-travel angle before starting the forward movement of the "Helipot".

A limit switch at the far end of the positioning device prevents the sliding contact of the "Helipot" from overshooting its limiting position at counting rates lower than $n_{\min}$.

Besides controlling the counting operation and resetting, starting and stopping the "Helipot" at the correct moments with respect to the beginning and the end of a counting interval, the computer performs several other operations after the counting is completed: it feeds the final "Helipot" current to the recorder, at the same time it energizes the chart drive of the recorder to advance the chart a fixed increment and it advances the detector tube arm of the goniometer a preset 2Θ-increment (cf. I). It then switches off the "Helipot" current and the recorder chart motor and energizes the reset solenoid. All these switching operations are controlled by means of a multiple-deck motor-driven cam switch (cycling switch), which can be seen in *fig. 4* and which completes one cycle in 60 sec (recycling time, which is the lower limit for the duration of a single measurement).



86162

Fig. 4. Computer chassis. A number of elements shown in fig. 3 are visible; the lettering is identical to that of fig. 3. To the extreme left is a multiple-deck motor-driven cam switch S which controls each cycle of the operation of the computer.

Reproducibility of measurements

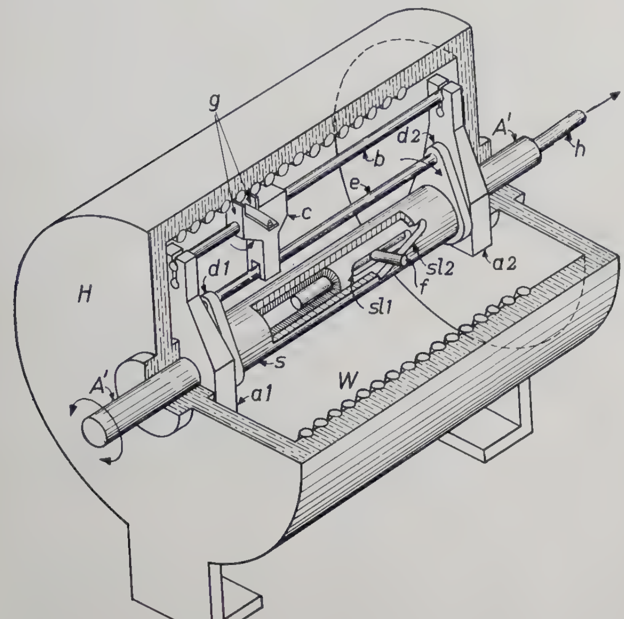
A number of precautions to ensure reproducible measurements have already been mentioned in the preceding section. They are aimed at maintaining the exact value of the pre-travel angle α and at making the time intervals t in eq. (1) and (3) strictly identical.

To conform with the latter condition, reliable operation of the friction clutch assembly, with a minimum and constant time lag between the energizing of the solenoids and the movement of shaft A, is of crucial importance. This was obtained by a careful construction of the linkage between the solenoids and the clutch faces. These faces are simply formed by the sides of the gears facing the drum. During axial movement of the gears, the teeth of the gears act as splines. Solenoid armatures and other moving parts are made as light as possible. The drum is faced on each end with a "Neoprene"-cork composition to assure positive driving action and a very rapid release. For the same reason the clearance is reduced to about $^1/_{64}''$ (0.4 mm); the motion is little more than enough to release the pressure. The shaft A on which the drum is fastened must, of course, have no noticeable end play.

Backlash in the gears should be very small, lest the slight shock from the solenoids should cause erratic operation. On the other hand, the gears must not be meshed too tightly causing them to move non-uniformly. A total backlash of about 0.12° is permitted, since the helical slide-wire is made up of a coiled resistance wire, so that the resistance changes in small but finite steps; each turn of the wire (having a resistance of 3 ohms) corresponds to about 0.003″, which, on a 3″ helix diameter, represents 0.12°.

Another important design detail affecting the reproducibility of the measurements is the sliding contact of the "Helipot". Although the contact and brush assembly normally supplied in the "Helipot" is satisfactory for many purposes, it was found to be too erratic for our application, especially in recording counting rates near the top of the selected range. In fact, our value of $R_{\min}$, the fixed resistance in series with the slide-wire, is 869 ohms and the 15 turns of the slide-wire have a resistance of 6780 ohms each, so that the first turn of the helix will cover the large range from $n_{\max}$ ($= Nk/869$, eq. 4) to 12.8% of $n_{\max}$ (viz. $Nk/(869 + 6780)$). In this large range one simple coil step (3 ohms) of the slide-wire near the top corresponds to a change of 0.34%. It is therefore important to record the current at the exact position of the sliding contact, which should not be off by more than one coil step, i.e.

0.003″ (0.08 mm). Now, the contact will pass the first part of the slide-wire twice in every single measurement so that excessive wear of this part, which should have the most precisely constant resistance value per unit of length, is a real danger.



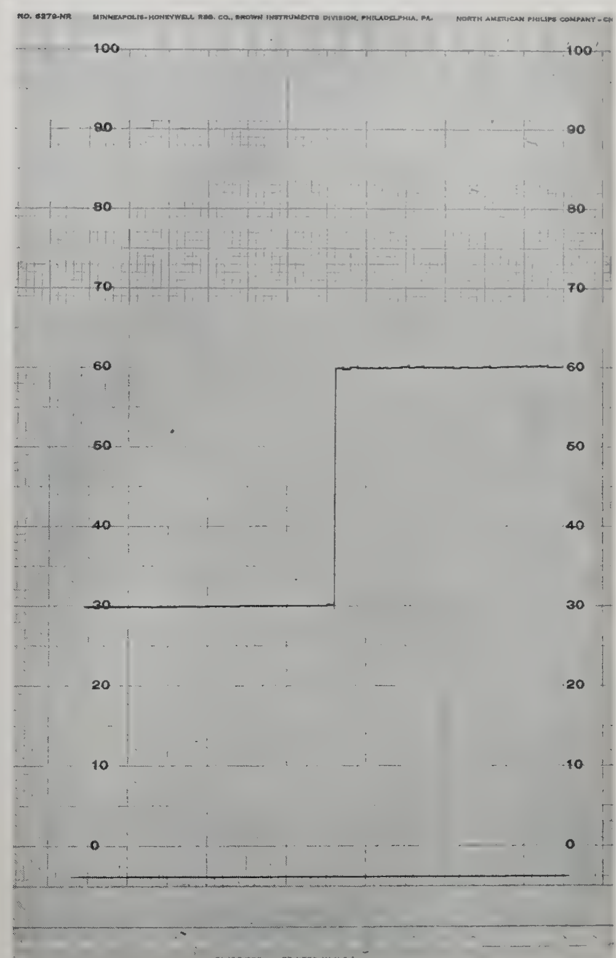
86458

Fig. 5. Retractable contact assembly developed for the "Helipot". The helical slide-wire W is secured to the inner wall of stationary cylinder H. The hollow shaft A' rotating within this cylinder, carrying the arms a_1 and a_2 and tie-rod b, is entrained by the main shaft A of the computer as indicated in fig. 3. The "Bakelite" guides g, gliding between adjacent windings of the slide-wire, cause the contact brush holder c to traverse rod b and thus to describe a helical path parallel to the winding. The contact will touch the wire, however, only when the contact holder is slightly rotated on the tie-rod in the direction of the arrow. Such a rotation can be brought about by a slight rotation of sleeve s carrying tie-rod e secured by arms d_1 and d_2. Shaft A' and sleeve s each have a slot, sl_1 and sl_2 respectively. A pin f engaging both slots locks shaft and sleeve so that they rotate together with a certain "phase angle" between the rods b and e which depends on the position of the pin, since slot sl_2 is inclined with respect to sl_1. The pin f is carried by the auxiliary shaft h, rotating with shaft A' and sleeve s. The relative angular position of tie-rods b and e is thus controlled by the axial position of the shaft h; a movement in the direction of the arrow causes the necessary relative rotation of sleeve s to lower the contact onto the slide-wire. During such a movement, shaft A' remains stationary owing to the friction brake B of fig. 3. The axial movement of shaft h is effected by a small electric motor (M_1 in fig. 3) controlled by the cycling switch mentioned in the text (S in fig. 4).

For these reasons the contact shoe and brush assembly of the "Helipot were modified in such a manner that the contact is lifted away from the wire whenever the shoe is in motion [4]). When it is required to record the result of a measurement the contact is brought down on the wire automatically but only just long enough to complete the recording. A cut-away view of the "Helipot" showing the inside construction with this retractable contact mechanism is given in *fig. 5*. The contact is lifted by means of a small motor (M, in fig. 3) which is controlled by the cycling switch and provides smooth reproducible action. The friction brake mentioned earlier (B in fig. 3) keeps the "Helipot" shaft from rotating while the contact is being actuated.

Performance

The over-all degree of reproducibility obtained is illustrated by *fig. 6*. The right hand part shows a recording made with the power line frequency of 60 c/s with rate k and scale factor chosen to make the full scale 100 counts/sec. There are a few steps in the line but it should be remembered that the reproducibility of the scaling circuit and associated



86164

Fig. 6. In order to demonstrate the degree of reproducibility obtained, a recording of a constant counting rate (power line frequency of 60 c/s) is made. At a full scale range of 100 counts/sec (right-hand line), the recording shows a few steps. At a larger full scale range, say of 200 counts/sec (left-hand line), no steps appear, owing to the fact that for a counting rate of 60 c/s a larger part of the slide-wire is now used, so that the discrete turns of the coiled slide-wire, which determine the smallest detectable difference in counting rate, have less effect.

---

[4]) E. A. Hamacher and K. Lowitzsch, U.S. Patent No. 2 658 131, Nov. 3, 1953.

circuits and relays are also included in the test. Because of the smaller relative influence of the discrete turns of the coiled slide-wire at the lower portions of the scale (see above), the reproducibility is even better there: no steps appear in a recording of the 60 c/s frequency at 200 counts/sec full scale (fig. 6, left).

It will be evident from the section on the principle of the computer that deviations from the desired linear relationship between the recorded current and the counting rate will occur when the fixed resistance $R_{\min}$ and the pre-travel Δ (or pre-travel angle α) are not adjusted exactly to their (interrelated) theoretical values. Another cause of non-linearity is due to R_l, the resistance of the slide-wire per unit length, not being strictly uniform throughout. The specified tolerance for the commercially available 15 turn "Helipot" is 0.1%, i.e. 100 ohms out of a total of about 100 000 ohms. If this total deviation were to occur in the first turn, covering the range from 12% to full scale, a maximum departure of 1.5% from linearity would be found in the response of the computer in this range. Actual experience shows that the maximum departure is always less than 1%. Because of the importance of the first turn of the "Helipot", the average value of R_l in this part should be used for the calculation of the pre-travel $\Delta = R_{\min}/R_l$.

Application of the Cooke-Yarborough method to the computer

The fixed count method will require very long counting times for low counting rates. Although this is a logical consequence of the condition of constant statistical relative error for all intensities, a compromise between accuracy and time consumption will often be desirable. As was explained in article II, a good compromise can be obtained by adding pulses of a *constant* controlled rate, produced by a separate pulse-generator, to the random pulses coming from the detector tube (Cooke-Yarborough method). Provisions have been made for using this method with the counting-rate computer. The changes required are quite simple and may be outlined as follows:

Summarizing the basic idea of the computer we may say that a counting rate n will give rise to a proportional current I; $n_{\max}$ and $I_{\max}$ constitute the limiting case. Owing to this proportionality, a constant counting rate f added to n will give rise to a constant additional current $I_f = If/n = I_{\max} f/n_{\max}$. If, therefore, we subtract this constant current I_f from $I_{\max}$, the remaining current will again be proportional to n and can be recorded as such. The only difference from the normal procedure will be that the uppermost part of the scale on the chart, viz. from $n_{\max}$ down to $n_{\max}(1 - f/n_{\max})$, is not used.

In order to make the used part of the scale the same for all selected ranges $n_{\max}$ (giving an identical calibration of 0-100 for this part for all ranges), it is necessary to make $f/n_{\max}$ a constant fraction. In the present instrument this fraction is made equal to 0.075: this permits the generator of the added pulses to be synchronized with the power line for $n_{\max}$ = 800 counts/sec, since f will then be equal to 60 c/s. For other possible values of $n_{\max}$, f will be equal to harmonics or sub-harmonics of 60 c/s. The gain in time for low counting rates is quite considerable with these added pulse rates, while the statistical relative error is not prohibitively increased. This was illustrated by fig. 15 and 16 in article II.

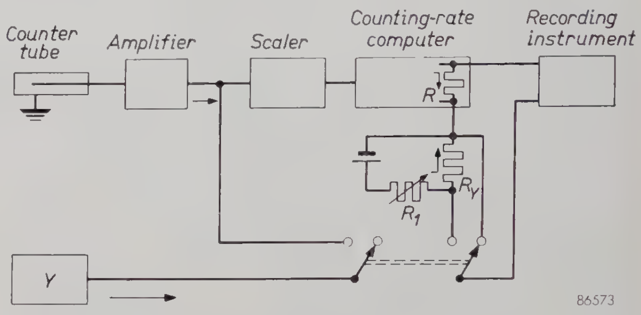


Fig. 7. Block diagram showing the provisions made for the application of the Cooke-Yarborough method to the computer. The pulse generator Y is connected to the input of the scaling circuits in parallel to the detector. At the same time a resistor R carrying a constant current (adjusted once and for all by means of a rheostat R_1) is connected in series with the resistor R through which the output current of the computer is fed. The voltage across $R + R_Y$ is fed to the recorder.

The conversion from normal operation of the computer to Cooke-Yarborough operation involves simply throwing a switch. The output of the pulse generator is thereby connected to the negative input terminal of the scaling circuits, in parallel with the pulses from the detector tube, and at the same time a circuit effecting the subtraction of current $I_f = 0.075 \times I_{\max}$ from the computer output current is switched on. This is shown in the block diagram *fig. 7*.

Summary. In measuring radiation intensities by counting quanta, the statistical relative error is reduced to a predetermined value by accumulating a fixed number of counts, N. The necessary time t is measured and the counting rate (which is proportional to the intensity) computed from $n = N/t$. Application of this method with the "Norelco" X-ray diffractometer for measuring diffracted intensities at narrow-spaced diffraction angle positions (which may involve many hundreds of measurements) is bound to make excessive demands in the operator's time, unless the process can be made automatic. This has been achieved by the development of a computer, evaluating n from the above equation by means of an analogue and recording it on a strip chart with a linear scale. The analogue used is Ohm's law: the contact of a slide-wire is moved at a constant rate for the time t necessary to accumulate a pre-set number of counts N; at the end of this time, a fixed voltage is applied across the slide-wire and the current which flows is fed to the recorder. The range can be adjusted by changing the speed of movement of the contact. In the mechanical design, many precautions were necessary to ensure reproducible operation: the double clutch drive for operating and resetting the sliding contact is designed to start and stop the movement at the exact moments required; a device is incorporated for resetting the contact to its precise starting position; in addition, a special retracting mechanism for the contact is incorporated to prevent excessive wear of the slide-wire. Reproducibility to within one turn of the coiled slide-wire is obtained, which corresponds to an accuracy of 0.34% at the upper end and much better accuracy at the middle and lower parts of the scale. Provisions are also made for applying the Cooke-Yarborough modification of the fixed count method with the "Norelco" computer.

AN ELECTRICAL CLINICAL THERMOMETER

621.317.39:536.531:616-073.65

Mercury clinical thermometers require a period of the order of 3 to 10 minutes before a reliable reading can be taken. There are types available of reduced size, which permit the reading time to be brought down to 1 minute. This reduction in size, however, makes reading as well as "shaking down" more difficult.

The long reading time is caused not only by the high thermal resistance between the body tissue and the mercury, but also by the considerable heat capacity of the thermometer. This heat capacity causes the tissue in contact with the thermometer to drop in temperature. It takes a considerable time for the temperature to be restored by the blood circulation.

The above disadvantages are almost entirely overcome in the electrical resistance thermometer [1]) shown in *fig. 1*, which has a reading time of only 13 seconds. The thermometer consists of a probe connected by a flex to a small meter easily held in the hand. The temperature-sensitive element in the probe is a ceramic semi-conducting material with a high negative temperature coefficient of resistance [2]) (NTC resistor).

The body of the resistor is oval in shape; the longitudinal axis is approximately 1 mm long and the diameter is approximately 0.4 mm. Two connecting wires are baked into the body of the resistance parallel to the longitudinal axis. The resistance measured between these two wires is a measure of the temperature. The outside of the NTC resistor is glazed to protect it from the influence of the atmosphere; this makes these resistors very stable, a matter of primary importance for the end in view. By a careful choice of composition and treatment of this ceramic material, its resistance at room temperature can be varied within wide limits.

Fig. 1. Left, the electrical clinical thermometer, consisting of a probe and an indicating instrument, connected by a flex. The button visible on the right of the instrument is pressed in order to take a reading. The reading may be taken about 13 seconds after insertion of the probe. Right, the model for skin thermometry.

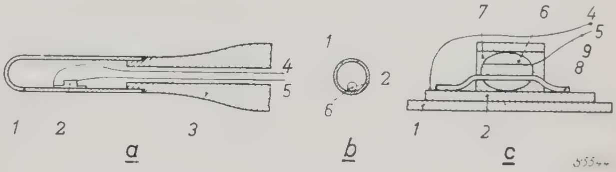


Fig. 2. *a*) Longitudinal cross-section of the probe of the clinical thermometer. *1* silver "bulb", a closed tube of the same form and dimensions as the bulb of a mercury clinical thermometer. *2* small silver cylinder soldered to the interior of the bulb and containing the NTC resistor. *3* hollow holder of nylon. *4* and *5* connecting wires.
b) Transverse cross-section of the bulb at the position of the small cylinder *2*. *6* is the NTC resistor.
c) Longitudinal cross-section of the silver cylinder *2*. *6* NTC resistor, *7* "Araldite" bonding resin. *8* thick wire of pure platinum, both ends being soldered to the lips provided on the cylinder *2*. *9* thin wire of platinum-iridium.

The end of the probe (*fig. 2*) consists of a small silver tube whose external dimensions are the same as those of the mercury bulb of a normal clinical thermometer. A small silver cylinder is soldered against the inner wall of this "bulb" and the closely fitting NTC resistor is cemented in it with "Araldite". There is no electrical contact between the cylinder and the outside of the NTC resistor but there is a good thermal contact; heat absorbed through the "bulb" from the body tissue is thus transferred to the NTC resistor via the circumference of the inner silver cylinder. In addition to this, the heat can flow directly into the interior of the NTC resistor via one of the wires (*8* in fig. 2*c*), which is therefore made comparatively thick (100 μ) and of pure platinum (like silver, an excellent conductor of heat). Both ends of this wire are soldered to the small lips on the cylinder surrounding the NTC resistor.

Clearly a current must flow through the NTC resistor in order that the measurement may be carried out. Owing to the good thermal contact between the NTC resistor and the silver "bulb", the heat generated in the former by this measuring current is rapidly distributed over the latter. This is most important: the heat generated causes a

---

1) The thermometers discussed in this article are at present used only for experimental purposes and are not generally available.
2) E. J. W. Verwey, P. W. Haayman and F. C. Romeyn, Semi-conductors with a high negative temperature coefficient of resistance, Philips tech. Rev. **9**, 239-248, 1947/48.

rise in temperature in the NTC resistor and consequently an error in measurement, which may be considerable unless the heat is rapidly dispersed. Tests have shown that in maintaining this error below an acceptable value, more heat may be dissipated in the NTC resistor with a thermometer of this construction then when the resistor is encased in a glass tube. This means that a higher current may be used, which makes the instrument more sensitive. The maximum permissible energy dissipation in the NTC resistor with the present design is 0.25 mW.

In order that the temperature of the object being measured shall be affected as little as possible by the measurement, conduction of heat from the object via the probe to the surroundings must be as small as possible. With this in mind, the holder to which the silver bulb is attached is made of nylon, which is a good heat insulator (much better than glass, for example). In addition, the surface of the bulb in contact with the nylon holder is made small with respect to the surface in contact with the body. Loss of heat via the copper connecting wires which link the NTC resistor to the meter is kept low by making these wires very thin. The error caused by heat loss via the probe and the error caused by heat generation in the NTC resistor as a result of the measuring current are opposed to each other and tend to cancel each other.

Yet a third source of error should be taken into consideration: this is operative even when the total dissipation of heat via the probe is so small that it does not noticeably affect the body temperature being measured. The silver bulb then has exactly the right temperature. Now the heat conducted away by the connecting lead *5* (fig. 2*c*) has to pass through the NTC resistor itself. Owing to the considerable thermal resistance of the latter, a temperature drop is set up across it, which leads to an error in measurement. This error can be reduced by increasing the thermal resistance of the conduction path beyond the NTC resistor. With this in mind, the wire *9* is very thin and is made of platinum-iridium whose thermal conductivity is about $^1/_3$ that of pure platinum. When these measures are taken, the heat resistance per cm length of the thin wire *9* is about 25 times as great as that of the thick wire *8*. Heat conduction along connecting wire *4* which is soldered to the small silver cylinder *2* (and thus electrically connected to wire *8*) and heat losses through the holder both bypass the NTC resistor, and do not therefore contribute errors of this nature.

The NTC resistor forms one of the branches of a Wheatstone bridge. The connecting leads run through the hollow holder and are joined via a flex to the measuring instrument which contains the other three resistance branches, the indicating meter and the supply battery. The bridge is not balanced for each reading, i.e. the instrument is not operated on the null principle: this would take up too much time. Instead, the resistances are so chosen that there is no current flowing through the meter (i.e. the bridge is in balance) when the temperature of the NTC resistor corresponds to the lowest value shown on the scale (35 °C). At higher temperatures the meter shows a deflection due to the out-of-balance current then flowing. The current flowing at temperratures lower than 35 °C causes the needle to deflect against a stop.

The battery is connected only during a measurement, by depressing a spring-button (fig. 1). Since the instrument can conveniently be read with the probe actually in position and almost immediately after its insertion, it is unnecessary to provide a maximum-reading device in the meter.

The meter is calibrated in °C, and its range runs from 35 to 42 °C, with a scale length of 3 cm.

The sensitivity of the meter circuit, i.e. the deflection of the meter per °C temperature variation in the NTC, depends upon the terminal voltage of the battery, which must necessarily be very stable in order to maintain a constant calibration. The mercury-oxide cell used meets this requirement provided that it is used only intermittently and briefly — as is the case for such temperature measurements — and provided that only a small current drain is involved [3]). The instrument includes a provision for a simple check on the battery voltage. With the voltage supplied by this type of cell (1.34 V), and the above-mentioned maximum permissible dissipation in the NTC resistor (0.25 mW) the values for the bridge-resistances are chosen so as to give maximum bridge sensitivity, i.e. maximum meter deflection per °C. Under these conditions, the NTC resistor has a value of the order of 1000 Ω and the total current drawn from the battery is small enough to be provided by a mercury-oxide cell of the smallest obtainable type. The indicating instrument is a moving-coil meter with a phosphor-bronze strip suspension; this gives a very rugged instrument which can withstand the shocks that a hand instrument may be expected to suffer. The temperature can be easily read to 0.1 °C which is the accuracy laid down in English and American standards.

In some cases it is important for the physician to be able to measure the temperature of the skin, e.g. in cases of disturbed blood-circulation conditions.

[3]) See, for example, G. W. Vinal, Primary batteries, J. Wiley, New York 1950.

For this application (measuring range 15 to 42 °C) it was necessary to change the form of the sensitive element to that shown in *fig. 3*. The skin thermometer is also shown in the photograph of fig. 1. With this larger measuring range, the fact that the resistance of the NTC resistor is not a linear function of the

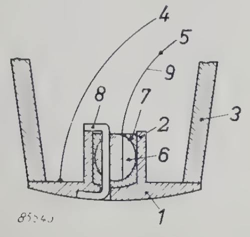


Fig. 3. Cross-section of the sensitive element of a skin thermometer. The element is designed so as to give good thermal contact with the skin. The numbers in the diagram have the same significance as in fig. 2. The components *1* and *2*, however, are now turned as a single piece from a rod of silver.

temperature becomes important; the resistance variation per degree temperature change decreases rapidly as the temperature rises (*fig. 4*). Fortunately, however, the sensitivity of the bridge for resistance variations of the NTC resistor increases with the increasing deviation from equilibrium of the bridge. A closer analysis of the properties of the bridge has shown that it is possible to design the bridge such that these two non-linear effects practically compensate each other, so that the relationship between the temperature and the bridge current is linear within an accuracy of 0.2 °C. The resistance values then used in the bridge for the skin thermometer (range 15-42 °C) happen at the same time to give maximum sensitivity.

A linear relationship offers considerable advantages. Not only is the absolute reading accuracy constant within the whole range but it is also possible to use a standard printed linear scale in the

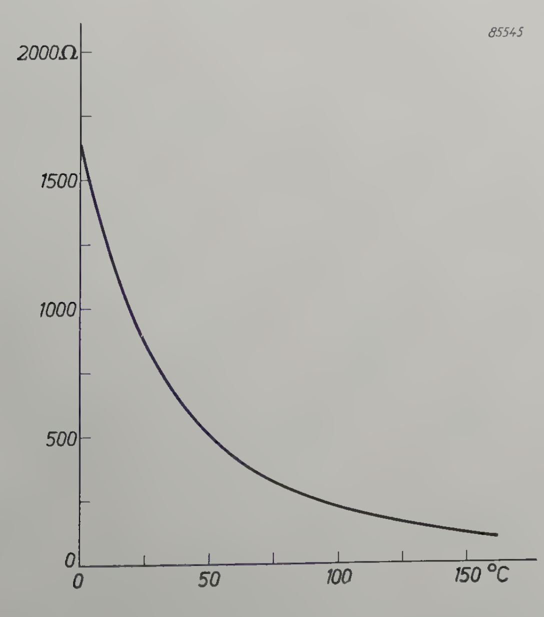


Fig. 4. The resistance of an NTC resistor as a function of the temperature.

meter. If the meter is adjusted to give correct readings at two points, it will then be correct over the whole scale. For one of these temperatures, viz. the lowest scale reading, the bridge is balanced by adjusting the resistance in one of the arms; for the second temperature, the sensitivity of the bridge can be adjusted by connecting a small resistance in series with the meter.

The skin thermometer described here may also be used as a clinical thermometer if used per rectum (if the thermometer is used e.g. in the arm-pit, the design of fig. 3 is not such as to always ensure the direct contact of the sensitive element with the skin). Thus, by providing the instrument with two ranges, selected by a switch, it is possible to combine the skin thermometer and the clinical thermometer in one instrument. This is a doubly attractive proposition since the construction shown in fig. *3* is simpler and thus cheaper than that shown in fig. 2.

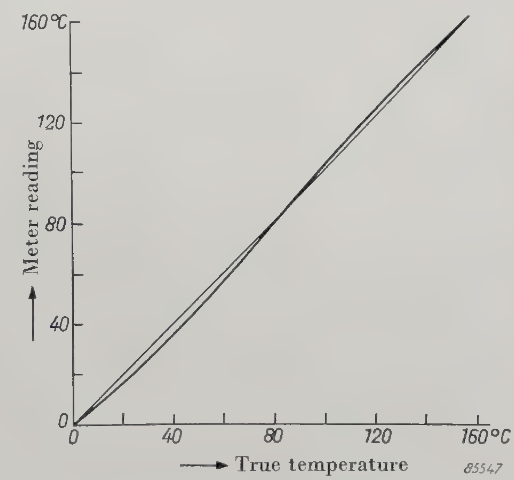


Fig. 5. Graph showing the slight deviation from a linear scale of an electrical thermometer with a range of 0-160° C. The temperature coefficient of the NTC resistor varies by a factor of 20 in this temperature range.

It is possible also to construct non-clinical thermometers with much greater ranges (e.g. 0 to 160 °C) based on the same principles. It is then no longer possible to keep the relationship between temperature and meter deflection exactly linear, so that a calibrated scale is necessary. It is possible, however, to choose the bridge resistances so that the deviation from linearity is reasonably small, so as to retain the advantage of an almost constant absolute reading accuracy (see *fig. 5*). Possible applications for such a thermometer include: rapid checking of the temperature of resistors in electrical apparatus during operation, and the temperatures of bearings, carbon brushes, transformer cores and electrode lead-in wires in electronic tubes.

C. C. J. ADDINK.

STEREO REVERBERATION

by R. VERMEULEN.

534.844:534.846.4:621.395.625.3

While too long a reverberation time makes speech unintelligible, a reverberation time which is too short makes music sound "dry" and brittle. Many varieties of acoustic materials are available for shortening the reverberation time and improving the intelligibility. Lengthening of the reverberation time and — perhaps more important — making the sound diffuse, can be achieved by electro-acoustic means. Tests have shown that in this way a good theatre hall can be made suitable for concerts.

Introduction

The acoustic properties of a hall are determined by the behaviour of the sound waves, in particular those which are reflected from the walls. The importance of the reflected sound can be well understood if one has listened to an open-air speech made without an amplifier installation, or if one imagines an open-air performance by a symphony orchestra without this aid.

A simple calculation will show that in a hall at even quite a small distance from the orchestra, the direct sound can sometimes be weaker than the reflected sound. At a distance r from a sound source of power P, the energy density of the direct sound is equal to $P/(4\pi r^2 c)$, where c is the velocity of sound. In a hall with a volume V and reverberation time T (defined as the time in which the sound intensity decreases by 60 dB, after the source has ceased radiating), the energy density of the indirect sound [1]) is $PT/(13.8\ V)$. These two energy densities are equal at a distance $r_0 = (1.1\ V/cT)^{\frac{1}{2}}$. Values of r_0 corresponding to various practical values of V and T are as follows:

$V =$	100	1000	10 000 m³
$T =$	0.7	1.0	1.5 s
$r_0 =$	0.7	1.8	4.5 m.

At distances greater than r_0 the indirect sound predominates. It is seen that this can occur at distances of only a few metres.

The qualities required of a hall for speech and for music are quite different. In a theatre, the intelligibility is of primary importance. If speech is to be clearly understood, the reflected sound must reach the audience with so little delay that it reinforces the direct sound but does not overlap the sounds which follow. For the latter, the persistence of the preceeding sound must be regarded as "background noise", which adversely affects intelligibility. As a rough guide, one can say that all sound which reaches the audience within 50 milliseconds can be regarded as useful sound [2]). Erwin Meyer has formulated the idea of "clearness" or "definition" [3]), which he defines as follows:

$$\frac{\int_0^{50\text{ msec}} p^2(t)\, dt}{\int_0^{t_1} p^2(t)\, dt},$$

where p is the sound pressure, t is the time (measured from the moment at which the source is silenced) and t_1 is a time much greater than 50 msec.

For a concert hall, on the other hand, the first requirement is not intelligibility, but a fine, full tone. Here it is much more difficult to specify the requirements. For speech, the reverberation must be accepted as an inevitable, disturbing accompaniment to the useful sound, because it simply is not possible to silence the sound suddenly after 50 msec. For music we know that the reverberation time not only may be, but must be, longer. The optimum value is clearly dependent on the nature of the music. It is often seen that a composer has consciously taken into account the acoustics of the space (church, concert hall, room) where he wanted his music to be played. The inclination to sing in the bathroom can probably be largely attributed to the long reverberation time of this acoustically "hard" room.

It is becoming increasingly clear that the reverberation time is not the only property governing the suitability of a hall for musical performances.

[1]) See for example A. Th. van Urk, Auditorium acoustics and reverberation, Philips tech. Rev. **3**, 65-73, 1938.

[2]) See the article referred to in [1]), pp. 72-73.

[3]) E. Meyer, Definition and diffusion in rooms, J. Acoust. Soc. Amer. **26**, 634, Sept. 1954.

One might even conjecture that here, too, reverberation is merely an inevitable subsidiary effect. Just as important, or perhaps even more so, is the "diffuse-ness" of the sound (and possibly also the nature of the fluctuations of the reverberation).

To study these phenomena more precisely, we attempted to produce an artificial diffuse reverberation in the laboratory by means of distributed loudspeakers which repeated the music played, with controllable intensity and lag. This experiment appeared to improve the acoustics to such an extent that we ventured to take the bold step of using this artificial reverberation to make a theatre suitable for concerts. We propose that this artificial diffuse reverberation be called "stereo reverberation".

Our first installation was in the Philips Theatre at Eindhoven, whose acoustical properties as a theatre were very satisfactory as a result of rebuilding in 1935, but which left much to be desired as a concert hall. In addition to this theatre, a hall known as the "Gebouw voor Kunsten en Wetenschappen" (Arts and Sciences Hall) in The Hague, is now fitted with a permanent installation for stereo reverberation [4]).

Principle of the installation for stereo reverberation

The delay wheel

The principle of the stereo reverberation installation may be explained with the help of *fig. 1*. Controllable time lags are obtained by means of magnetic recording and playback. Magnetic material such as is used for magnetic tape is coated on the rim of a wheel ("delay wheel") [5]). The music is recorded on this material via a microphone and a recording head. A number of play-back heads — in the final apparatus four, in experimental types six (*fig. 2*) or more (*fig. 3*) — are mounted around the circumference of the wheel and connected via separate channels to loudspeakers. These are installed in various places in the hall: on the ceiling, along the balustrade of the balcony, in a lighting cornice, in "dead" corners under the balcony, etc. Between the last play-back head and

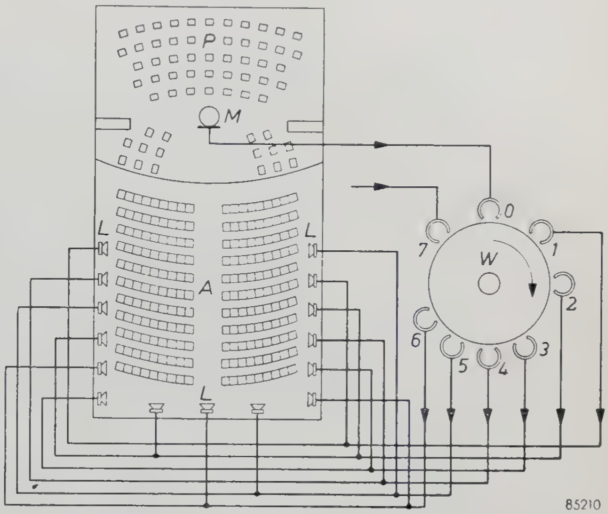


Fig. 1. Installation for simulating indirect sound with various time lags. *A* auditorium. *P* platform. *M* microphone. *W* delay wheel, coated on the edge with magnetic material suitable for magnetic sound recording. *0* recording head, *1...6* play-back heads. *7* erasing head. *L* loudspeakers.

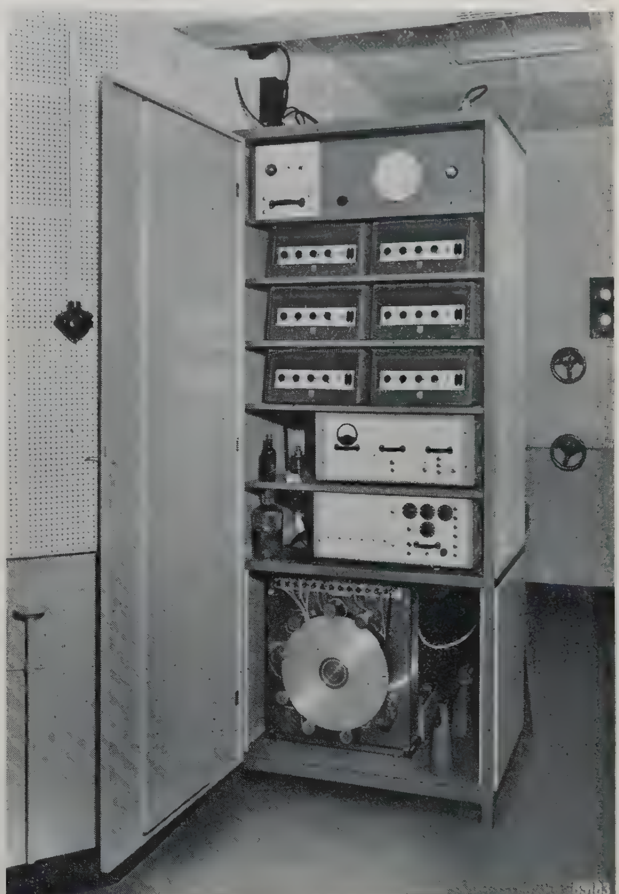


Fig. 2. Stereo reverberation installation in the Philips Theatre at Eindhoven. At the bottom of the cabinet is the delay wheel (cf. fig. 1); above it the amplifiers.

---

4) Demonstrations of stereo reverberation have also been given at the first I.C.A. Congress on Electro-acoustics, (Acustica **4**, 301, 1954), at Gravesano in Switzerland (at the invitation of the conductor W. Scherchen; see his book Musik, Raumgestaltung und Elektroakustik, Arsviva Verlag, Mainz 1955) and at the 3rd "Tonmeistertagung" of the Nordwestdeutsche Musikakademie Detmold (see D. Kleis, Elektron. Rdsch. **9**, 64-68, 1955 (No. 2)). Similar tests, but done in the open air, are described by H. S. Knowles, Acustica **4**, 80-82, 1954 (No. 1).

5) Others have also constructed a similar equipment, but have not created diffuse reverberation with it; see H. Schiesser, Einrichtungen zur Erzeugung künstlichen Nachhalls, Funk und Ton **8**, 361-368, 1954 (No. 7), and P. Axon and co-workers, Artificial reverberation, J. Instn. El. Engrs. **1**, 368-371, 1955 (No. 6).

the recording head is an erasing head, which ensures that the magnetic layer is blank on returning to the recording head.

For three reasons, however, the listener hears more than these seven (1 + 6) reports: firstly because to each play-back head, several loud-

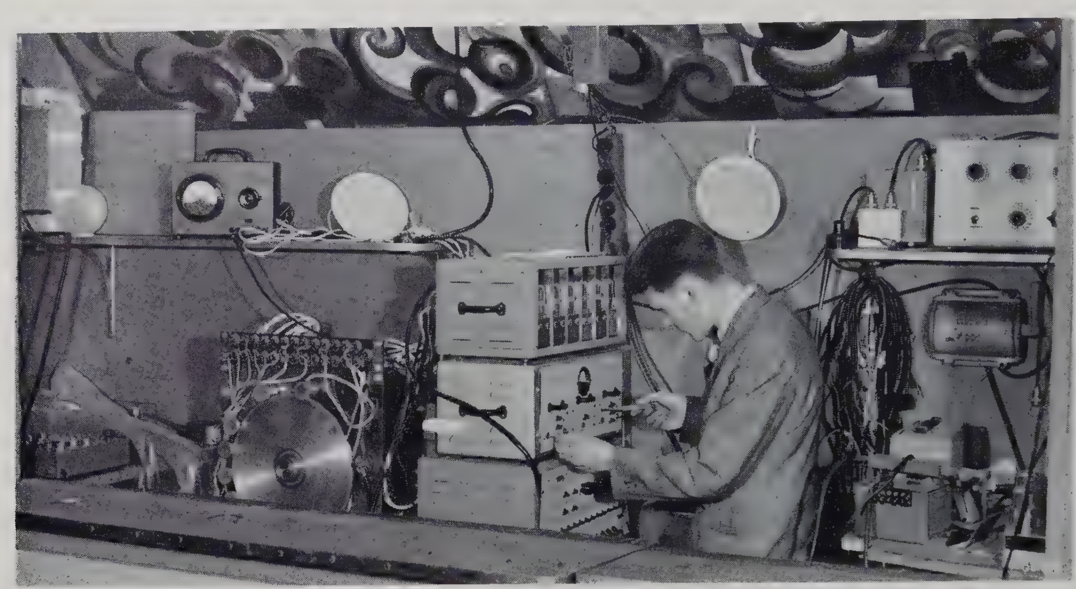


Fig. 3. Stereo reverberation installation in the "Gebouw voor Kunsten en Wetenschappen" (Arts and Sciences Hall), The Hague, in experimental form. The delay wheel was then fitted with ten play-back heads, of which six could operate simultaneously. It was possible to switch over rapidly from one set of play-back heads to another.

speakers are connected, which are dispersed throughout the hall and are at different distances from the listener, so that the transit times are different; secondly because the walls reflect both the original sound and those coming from the loudspeakers, and thirdly because the report from each loudspeaker is picked up by the microphone, once more recorded on the wheel and, though attenuated, reproduced six-fold.

Let us consider the case of a sharp report produced in the hall (pulse I_0, *fig. 4*). After a certain transit time this reaches the microphone and then the artificial indirect sound begins. The latter, if we are using six play-back heads, consists in the first place of six successive reports from the loudspeakers ($I_1 \dots I_6$). The intensity of each of these reports can be adjusted at will by the gain controls; the time intervals are also under control, by the spacing of the heads round the wheel and by its speed of revolution.

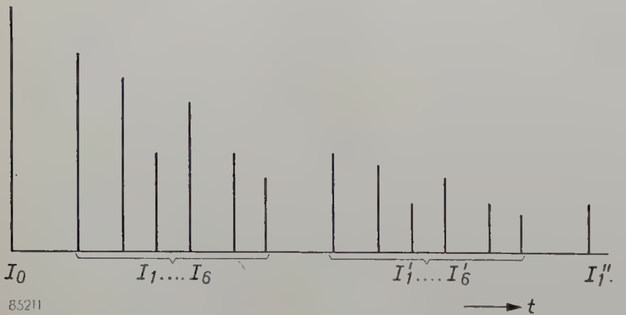


Fig. 4. If a report (pulse I_0) is produced in a hall with stereo reverberation, the six loudspeakers deliver during one revolution of the delay wheel the six reports $I_1 \dots I_6$. If electrical feedback is applied from the sixth play-back head to the recording head, the loudspeakers give a second series of six reports ($I_1' \dots I_6'$), a third, and so on. (Acoustic feedback is here neglected.)

All these effects contribute to the fact that the original report is followed by a large number of others, so that the reverberation time can achieve considerable values. The third effect mentioned above — the feedback from the loudspeakers to the microphone — even involves the danger that one particular note continues to sound too long and in the extreme case howls back continuously; therefore this effect must be curtailed as much as possible (we shall return to this point presently). This acoustic feedback can be advantageously replaced by electrical feedback, in which a chosen fraction of the signal from the last play-back head is fed back to the recording head. This records the signal afresh: in our example, a second series of six reports then ensues ($I_1' \dots I_6'$ in fig. 4); this series is once more recorded and a third series follows, and

so on. By adjusting the fraction of the output signal which is fed back, each series can be attenuated to any given degree, so that the reverberation time can be chosen at will.

Calculation of the effect of the stereo reverberation on the acoustic properties of the hall

The following calculation shows, for a simple case, the effect which stereo reverberation has on the acoustic properties of the hall. Suppose that the energy density in the hall is $E(t)$. Then $-V\mathrm{d}E/\mathrm{d}t$ is the rate at which acoustic energy in the hall diminishes. In a hall without stereo reverberation, this must be equal to the power absorbed by the walls [6]), which is proportional to E, say equal to αE. Stereo reverberation supplies extra power proportional to the energy density at some time τ ago: $\beta E(t-\tau)$. We can thus construct the equation:

$$-V\frac{\mathrm{d}E}{\mathrm{d}t} = \alpha E(t) - \beta E(t-\tau).$$

A solution to this is:

$$E = E_0 \exp(-mt),$$

where

$$mV = \alpha - \beta \exp(m\tau);$$

the quantitity m is inversely proportional to the reverberation time T, and equal to $13.8/T$. For small values of $m\tau$, an approximate value of $\exp(m\tau)$ is given by $1 + m\tau$. Thus we have:

$$m \approx \frac{\alpha - \beta}{V + \beta\tau}.$$

From this it is seen that increasing the strength of the stereo reverberation (β) has the same effect as decreasing the absorption (α) or increasing the volume (V) of the hall. An increased lag τ also gives the effect of a hall of larger volume; we shall return later to this point.

The microphones

The above calculation was based on the assumption that one could speak of "the" energy density in the hall. In many cases the total energy (sum of potential and kinetic energy) is indeed fairly evenly distributed throughout the hall. With standing waves, however, as is well known, the potential and kinetic energy alternate with each other, and this means that the sound pressure at constant frequency changes sharply from place to place, and, conversely, that at a particular place the sound pressure is very dependent on the frequency. The microphone which picks up the signal which is fed back to the hall via the delay wheel and the loudspeakers, responds only to the pressure at the spot. The factor β in our calculation therefore varies sharply with the frequency, and the reverberation time, which is inversely proportional to $\alpha - \beta$, varies also and to a much greater degree. If β as a function of the frequency shows a peak, then increasing the amplification will cause the note at which the peak occurs to go on sounding for a long time, while for most other frequencies, the reverberation time is not yet lengthened appreciably. With even greater amplification, the note fails to decay at all (howl-back: $\alpha - \beta$ has become negative).

The obvious solution is to try and suppress the peak by a filter in the microphone channel. However, there are so many peaks in the frequency characteristic of a hall, and these peaks are so sharp, that the suppression of all of them would be a hopeless task, particularly since the peaks are modified by all changes made in the hall or on the platform. It is worthwhile, nevertheless, to attenuate those frequency ranges in which the highest peaks occur in such a way that increase of amplification causes a number of widely separated frequencies to howl back at the same time.

In this connection it can be argued that electrical feedback via the delay wheel is a better means of obtaining a long reverberation time than acoustic feedback via the hall.

In a system with delayed feedback, the feedback signal and the input signal exhibit, in general, a relative phase difference which is proportional to the product of frequency and time lag. If there are a great number of feedback channels (as there invariably are in practice, with acoustical feedback), and if we suppose that they all make equal contributions to the input signal, then in general the contributions will show fairly random phase differences, so that the average resultant increases as the root of the number of feedback channels. There will be, however, one or more frequencies for which all the contributions are nearly in phase with each other. For these frequencies the resultant will be a maximum and equal to the algebraic sum of the contributions, and thus proportional to the number of feedback channels. This reasoning makes it clear that the ratio of the maximum value of the resultant (which occurs at a particular frequency and limits the maximum amplification) to the average value (at other frequencies) increases with the root of the number of feedback channels. In the case of a microphone and loudspeakers in the same hall, this number is very large and it is to be expected that a situation can easily arise in which one note continues to sound for a long time. With feedback via the delay wheel,

[6]) See for example the article referred to in [1]), where it is shown that the proportionality factor α is equal to $\frac{1}{4}aAc$, where a is the average absorption coefficient and A the area of the walls.

we are using only one feedback channel (from the last playback head to the recording head) and the above-mentioned danger is thus much less. It would be possible to introduce more feedback channels, e.g. from one or more of the preceeding playback heads to the recording head. Experiments have shown that this is undesirable, and this is partly explained by the above considerations.

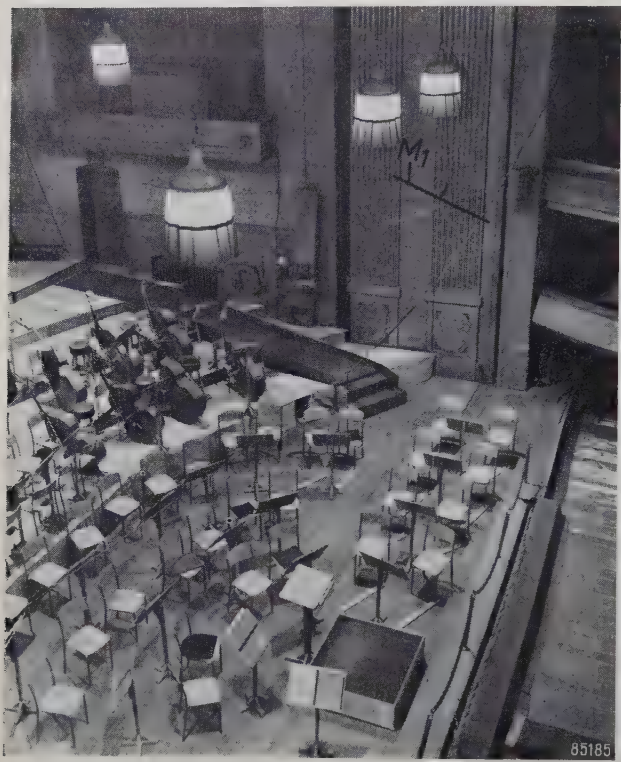


Fig. 5. The stage in the "Gebouw voor Kunsten en Wetenschappen" in The Hague. M_1 is the line microphone (rod with ten condenser microphones). (Other microphones visible in the photograph were for broadcasting and had nothing to do with stereo reverberation.)

Instead of a microphone which responds to the average sound density in the hall, the other extreme would be one which responds exclusively to sound coming directly from the source and is insensitive to sound from the hall [7]). The danger of notes howling back would then be completely averted. This situation can be closely enough approximated to by using a microphone having a sharp directional characteristic. An arrangement for achieving this, a so-called "line microphone", consists of a group of ten condenser microphones mounted at equal intervals along a rod rather more than a metre in length (*fig. 5*). By suitably positioning this line microphone above the orchestra, we can ensure that the loudspeaker signal at the microphone, although still making an important contribution, no longer dominates. If it is not possible to cover the orchestra adequately with one line microphone, two may be used (*fig. 6*).

The use of one or more extra microphones can be desirable in order to strengthen weak instruments in the orchestra. Thus, for example, in the Philips Theatre at Eindhoven the organ was rather weak in comparison with the orchestra and choir in the annual performance of Bach's St Matthew Passion; it appeared to be an improvement when this instrument was boosted by a microphone of its own (*fig. 7*). With such cases in mind, the stereo reverberation system was provided with several microphone channels with mixing facilities. These should, however, be used with the utmost restraint, for the sound engineer must never encroach upon the conductors prerogative for the balance of the instruments in the orchestra.

Fig. 6. In the Philips Theatre at Eindhoven two line microphones are used (M_1, M_2). The loudspeakers are mounted in a concealed position behind the lighting cornice L (compare fig. 9).

[7]) This occurs, for example, in the case of music reproduced in a hall, when it is (or was) actually performed elsewhere. We return to this point again at the conclusion.

Fig. 7. The small organ in the Philips Theatre can be reinforced through an amplifier channel fed from a separate microphone M_3.

The loudspeakers

As already suggested, the diffuseness of the sound is perhaps even more important than the lengthening of the reverberation time. Diffuseness can be obtained by dispersing the loudspeakers over the hall (*fig. 8*) and connecting them to the various play-back heads. The wiring is simplified if all the loudspeakers belonging to one group (fed from the same play-back head) are connected in series. In the case of four play-back heads, one four-core cable is run around the hall, balconies, etc; where a loudspeaker is installed, the appropriate core is cut and the speaker connected in series (*fig. 9*).

The distribution of the loudspeakers between the various play-back heads should be done as randomly as possible. The only restriction is that the audience should never get the impression that the sound comes from the loudspeakers. We shall now try to explain further the general lines to follow in order to avoid this impression.

We shall use some results of the work of K. de Boer on stereophony [8]). In fact, we are here dealing

[8]) For a recapitulation of the principles of stereophony, with references to the literature, see R. Vermeulen, Philips tech. Rev. **17**, 171-177, 1955/56 (No. 6).

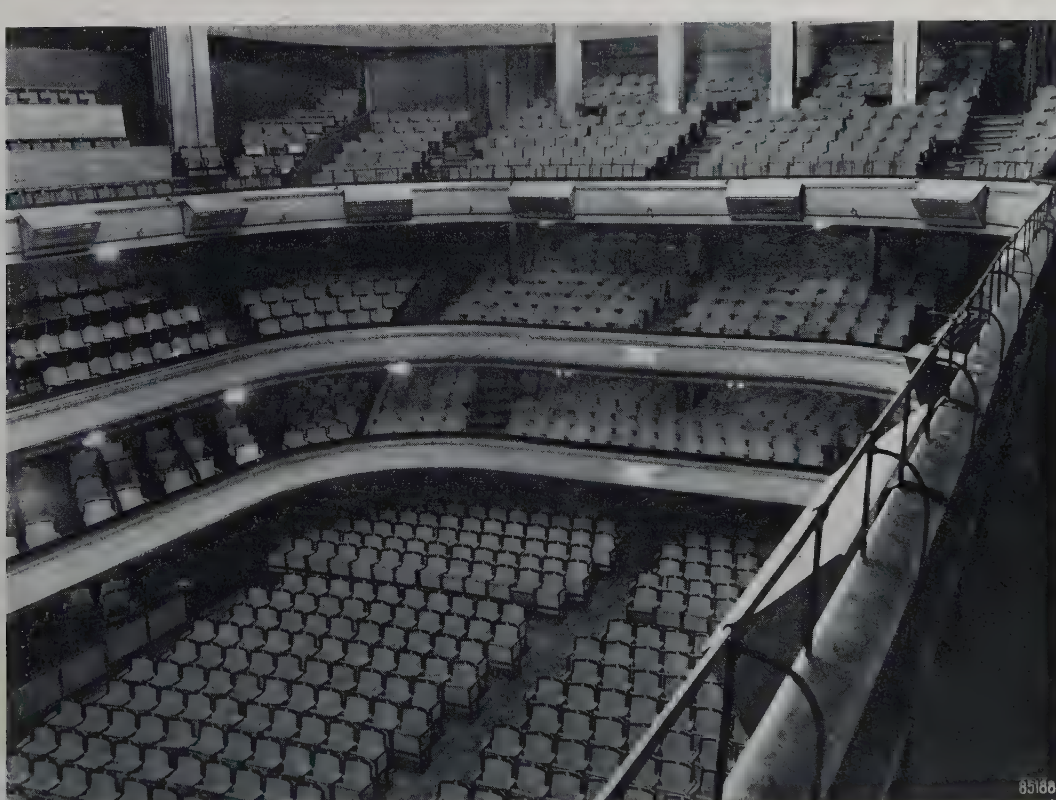


Fig. 8. In the "Gebouw voor Kunsten en Wetenschappen" the loudspeakers are mounted along the edge of the upper balcony and (not visible in the photograph) under the balconies.

with an analogous problem. The condition that nobody in the hall may consciously hear music coming from the loudspeakers, means that at all points in the hall the "sound image" must appear to be located in the orchestra. This latter can be

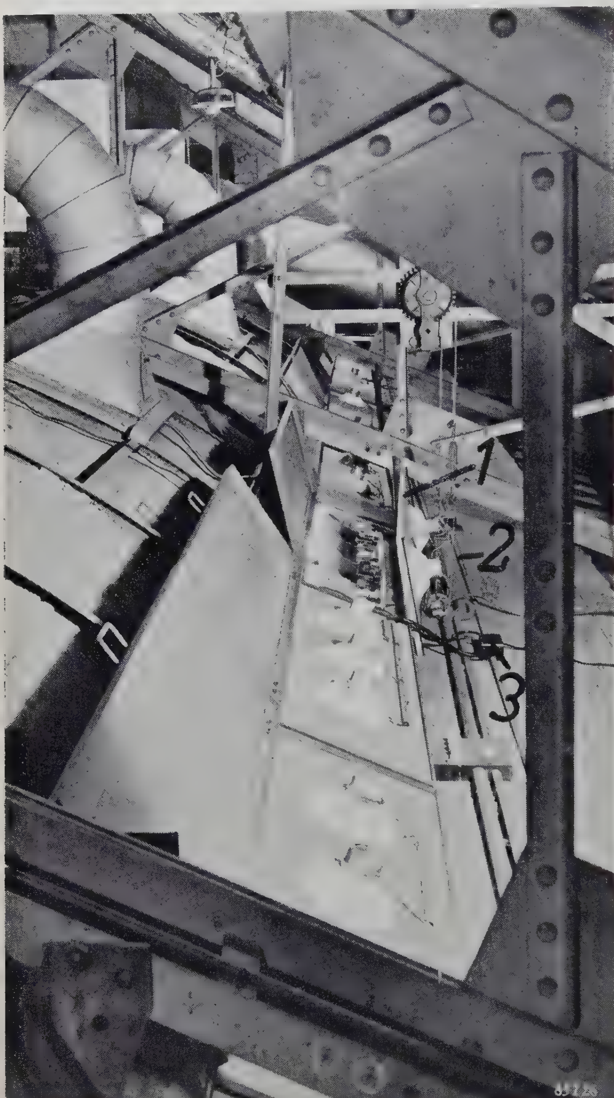


Fig. 9. The loudspeakers of the stereo reverberation installation in the Philips Theatre are mounted in groups of two or three in the panels which cover the lighting cornice (L in fig. 6). At *1* one of the panels has been lifted up. *2* is the (six-core) cable, *3* a junction box.

regarded as one of the two sound sources in a stereophonic installation, one of the loudspeakers being the other. For the sake of simplicity we assume that the observer is in the plane of symmetry of the two sources (if he is otherwise situated, the values to be mentioned presently must be modified by suitable amounts). We then know that he will locate the virtual sound source (the "sound image") in the orchestra if (a) the sound from the orchestra and that from the loudspeaker arrive at the same time but the intensity of the orchestra is at least 10 dB above that of the loudspeaker, or (b) orchestra and loudspeaker are equally loud but the sound from the orchestra arrives 2 msec earlier. These are the two extremes; for intermediate cases, as far as the position of the sound image is concerned, a lag of 1 msec in one sound is compensated by an increased intensity of 5 dB, according to the almost linear relationship [9]) shown in *fig. 10*. The sound image is still located in the orchestra even when the intensity of the orchestra is 5 dB less than that from the loudspeaker, provided the sound from the orchestra arrives 3 msec earlier.

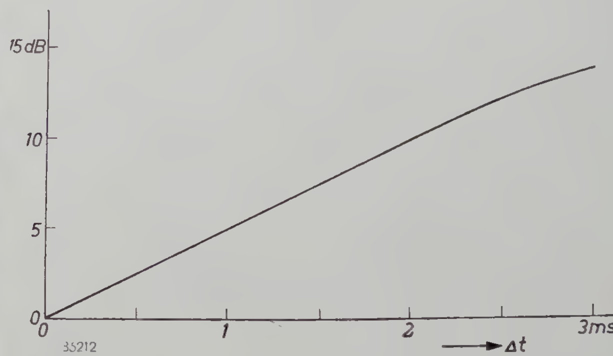


Fig. 10. Differences in intensity (in dB) plotted against the phase differences Δt which produce the same angular displacement of the sound image.

Another condition is that for no listener may the sound from the loudspeakers arrive with so great a time lag that it is no longer experienced as reverberation of the orchestra, but as a separate echo. This means that for no observer may the first loudspeaker signal which reaches him arrive more than 50 msec after the direct sound from the orchestra. This value corresponds to that found earlier when investigating the maximum time interval during which, in speech, the indirect sound contributes to the intelligibility (see Introduction). Recent investigations [10]) have confirmed this value and also that the indirect sound may be stronger than the direct without disturbing the location of the sound image.

On the basis of these data we can set down the following rules for a stereo reverberation installation:

[9]) K. de Boer, Stereophonic sound reproduction, Philips tech. Rev. **5**, 107-114, 1940.

[10]) H. Wallach, E. B. Newman and M. R. Rosenzweig, The precedence effect in sound localization, Amer. J. Psych. **62**, 315-336, 1949; G. Meyer and G. R. Schodder, Über den Einfluss von Schallrückwürfen auf Richtungslokalisation und Lautstärke bei Sprache, Nachr. Akad. Wiss. Göttingen, IIa, 31-42, 1952.

1) No member of the audience may receive sound from any loudspeaker before the direct sound has reached him.
2) Nowhere may the first loudspeaker sound arrive more than 40 msec after the direct sound (a safety margin of 10 msec has been deducted from the limiting value of 50 msec).
3) The intensity of the loudspeakers may nowhere be more than 5 dB above that of the direct sound [11]) (apart from the diffuseness required, this is another reason why many dispersed loudspeakers must be used).

If the same hall is also to be used for plays or lectures and improvement of the intelligibility is desirable, the stereo reverberation installation can also make itself useful for this purpose. One must then ensure that no sound is repeated later than 40 msec after the direct sound, and no feedback should be applied.

Directions for operation of the installation

Complying with the rules listed above does not necessarily ensure that the installation works satisfactorily. Because the theory lacks sufficient experimental backing, one should beware of clinging to unfounded preconceived opinions.

An example will illustrate this. In one of our first experiments with stereo reverberation, it was thought that the artificial reverberation should be built up of as many repetitions as possible, in order to obtain the smoothest possible exponentially decreasing intensity. It appeared, in fact, that though this did give the impression of a long reverberation time, this was by no means accompanied by the feeling of being in a large hall — rather that of a small "hard" room such as a bathroom. To suggest a large space, it was necessary to increase the time interval between the echos and to make the reverberation not at all smooth. By careful adjustment in a laboratory room of about 1000 m^3 volume, we could create the acoustic impression of being in a cathedral.

One general piece of advice relating to electro-acoustical intervention in musical performances is that the engineer must show considerable retraint. He must suppress his desire to make the effect *striking* and take care that he does not exaggerate. His is a thankless task: whenever his work is recognized for what it is, he will be reproached, and the better he does his work, the more natural the result will appear and the less thanks he can expect. The highest praise he can receive is probably the simple verdict: "...the orchestra sounded much better".

As we have said, not only the Philips Theatre at Eindhoven, but also the Gebouw voor Kunsten en Wetenschappen in The Hague, is fitted with a stereo reverberation installation. On November 30th, 1954, this installation (then still provisional) had its public debut during a concert given by the Residentieorkest. After the concert, the effects which can be achieved were demonstrated more emphatically.

The verdict on the operation during the concert varied from "favourable" to "very favourable". Some people, however, found it difficult to observe the effect consciously. This is illustrated by a remark from a musician in the audience: "It was remarkable that one could not consciously hear that the installation was in operation; one only felt, or experienced it. Only during the demonstration after the concert did I become consciously aware of it, and convinced that my observations during the concert were not imaginary".

Though the improvement may be appreciated only unconsciously by some of the audience, it is another matter for the performing musicians, both members of the orchestra and soloists; they experience the stereo reverberation very clearly and consciously as making the hall more "playable". This undoubtedly contributes to the attainment of a higher artistic level.

We feel that it is an important milestone in the development of electro-acoustics that leading musicians not only permit microphones and loudspeakers and all that goes with them, in the concert hall, but actually welcome their help.

Other applications of stereo reverberation

Stereo reverberation will undoubtedly find other applications apart from the conversion of theatres, acoustically speaking, into concert halls. However, the apparatus is so complicated that these applications will be limited to the professional field for the time being.

One obvious application is in broadcasting studios. Here, stereo reverberation can be a means of adjusting the acoustics of the studio to the nature of the music or the play to be performed. This can, of course, be done afterwards by adding artificial reverberation to the microphone signal; the echo chamber is a device often used for this purpose. But this deprives the musicians of the stimulus of good acoustics and if there is an audience in the studio, they too miss the full effect. Stereo

[11]) See fig. 8 in the article by Meyer and Schodder, referred to in [10]).

reverberation can overcome these difficulties [12]).

Another application that springs to mind is in the cinema. Here stereo reverberation can be obtained in the manner described above, with the difference that the microphone is discarded and the recording head on the delay wheel is now fed with the signal derived from the sound track of the film. It is even simpler if the cinema is fitted for a system like "CinemaScope", i.e. fitted with loudspeakers around the hall and projectors suitable for films with more than one sound track; the direct and the delayed diffuse sound could then be recorded beforehand on these tracks, so that the cinema has no need to be equipped with the delay wheel.

Finally we may mention the "duplication of concerts" [13]), that is the stereophonic reproduction in one or more "overspill" halls of a concert given elsewhere. Clearly, diffuse reverberation in the "overspill" halls can considerably increase the musical value of the programme presented.

[12]) According to a private communication from Dr. J. J. Geluk, Head of the Dutch Broadcasting Union Laboratories, the Union is planning to build two large broadcasting studios (7500 m^3 each) in which the principles of stereo reverberation will be applied.

[13]) R. Vermeulen, Philips tech. Rev. **10**, 169-177, 1948/49 and R. Vermeulen and W. K. Westmijze, Philips tech. Rev. **11**, 281-290, 1949/50.

Summary. A shortage of good concert halls means that music is often presented in a hall which is less suitable for this purpose, for example, in a theatre. Such halls can be given better acoustics for music by providing artificial diffuse reverberation. An installation is described with which such "stereo reverberation" can be provided electro-acoustically. It is based on a so-called "delay-wheel", the rim of which is coated with a material suitable for magnetic sound recording. A recording head records the music performed. A number of play-back heads (say, four or six) around the circumference of the rotating wheel pick up the recorded sound with predetermined time lags, and separately feed a group of loudspeakers, mounted throughout the hall. The intensity of each group is independently controllable, and the time lags can be regulated by the speed of the wheel. The sound is picked up by one or two "line microphones", each consisting of ten condenser microphones with strongly directional characteristics, to reduce the possibility of continuous "howl-back".

It has been shown that such an installation can produce a great improvement in the musical acoustics of a hall. Some of the audience experience this only unconsciously, but the performing musicians are strongly aware of it as making the hall more "playable".

Finally, some other possible applications of stereo reverberation are discussed: in broadcasting studios, in the cinema and in the "duplication" of concerts.

ABSTRACTS OF RECENT SCIENTIFIC PUBLICATIONS BY THE STAFF OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN

Reprints of these papers not marked with an asterisk * can be obtained free of charge upon application to Philips Electrical Ltd., Century House, Shaftesbury Avenue, London W.C. 2.

2217: A. Venema: Thermische emissie (T. Ned. Radiogenootschap **19**, 283-303, 1954). (Thermionic emission; in Dutch).

After an introduction to the subject, the author deals briefly with Schottky's method for calculating the work function, the influence of adsorbed layers on the work function, and its measurement on the basis of the Richardson formula. This is followed by a survey of emitter substances, with data on the temperature dependence of the current density. The thermionic efficiency is considered. The problem of the life of oxide cathodes is also dealt with. The shortcomings of the oxide cathode are mentioned, which have lead to the development of the so-called dispenser cathodes. Three types of the latter are discussed: the L-cathode, the impregnated cathode and the pressed cathode.

2218: D. Kleis: Dynamique de l'enregistrement magnetique (Onde électrique **34**, 753-760, 1954).

For good sound reproduction, an adequate dynamic range is essential. A large orchestra has a dynamic range of about 80 dB; with magnetic recording, a dynamic range of only 70 dB is attainable, so that it is important to ensure that this is used to the fullest extent. This puts high requirements on the recording apparatus as well as on the reproduction instrument. Unless special precautions are taken, there is the danger with such a large dynamic range that the noise level and distortion are unacceptably high; this is a result of the asymmetry of the erasing current but, more important, of the A.C. biassing current. With the usual type of play-back head, the signal voltage delivered to the amplifier is proportional not only to the amplitude of the flux in the play-back head but also to the frequency. This results in the signal voltage having a dynamic range considerably greater than that of the recording itself. To prevent the introduction of noise and distortion in the amplifier, special circuits must be used. This article discusses these factors and the means by which the desired dynamic range can be obtained in both recording and reproduction. The methods used include, in particular, the use of a negative feedback h.f. oscillator for erasing and

biassing during recording and an amplifier with frequency-dependent negative feedback during reproduction.

2219: G. W. Rathenau and G. Baas: Electron-optical observations of transformations in eutectoid steel (Acta Metallurg. **2**, 875-883, 1954).

More detailed account of investigations described in brief in Philips tech. Rev. **16**, 337-339, 1955/1956.

2220: J. M. Stevels: Networks in glasses and other polymers (Glass Ind. **35**, 657-662, 1954).

The dielectric losses, measured as tan δ, as a function of temperature are compared for a number of glasses, fused silica and crystalline quartz. It is shown that the sharp peak in the curve for quartz at low temperatures has its origin in a relaxation mechanism. For fused silica and the glasses examined, tan δ has a very broad maximum at low temperatures. It is reasonable to suppose that in the latter cases we are concerned with a movement of the oxygen, modified by the network in which it is bound. With decrease in Y (the average number of points of contact per tetrahedron), tan δ_{max} and also the corresponding decrease of the dielectric constant ε go through a maximum whereas the activation energy of the relaxation phenomenon goes through a minimum. This can be understood as due to the fact that although the network gets looser and looser, the network-modifying ions increase in number so that the oxygen becomes less mobile. Tan δ for silicones and other organic polymers also shows a broad maximum at very low temperatures. Here, however, sharp peaks are also present at about 200 °K. The possible origin of these peaks is discussed.

2221: F. A. Kröger and H. J. Vink: Physico-chemical properties of diatomic crystals in relation to the incorporation of foreign atoms with deviating valency (Physica **20**, 950-964, 1954).

If foreign ions of a valency deviating from that of ions of the base lattice are incorporated, electro-neutrality is maintained. This is effected in a number of ways, some of which are already known. For example, electro-neutrality may be maintained by the formation of vacancies, or by the occupation of interstitial sites. A second mechanism is that in which the incorporation of the foreign ions is accompanied by a reduction or an oxidation of the base lattice. It is shown that it is possible to consider the various ways of maintaining the electro-neutrality from a general point of view. There are a number of factors which govern the way the electro-neutrality can be maintained. Among these are the concentration of the foreign ions, the tendency of the base lattice to form lattice imperfections, the position of the energy levels associated with the various lattice imperfections, and the width of the forbidden zone. Another factor of importance is the atmosphere in equilibrium with the compound. In such a way, a third mechanism of maintaining the electro-neutrality is found. Application of these considerations to CdS as a base lattice gives a satisfactory agreement with experiment.

2222: F. van der Maesen and J. A. Brenkman: Acceptor activity of copper in N and P type germanium of different resistivity (Physica **20**, 1005-1007, 1954).

Experiments are carried out in which N and P type germanium samples with various resistivities or a bar with a resistivity gradient are saturated with copper at temperatures of 750 and 810 °C. The number of introduced acceptors calculated from the change of resistivity after quenching, is dependent on the position of the Fermi level. In material that remains N type after saturation, there is a considerable increase of the acceptor activity. It is therefore possible that copper produces extra acceptor levels with a rather high activation energy.

2223: H. J. G. Meyer: On the theory of transitions of **F**-centre electrons (Physica **20**, 1016-1020, 1954).

In the expressions for the probability of a radiationless transition in F centres, as given by Huang and Rhys and more recently by the present author, certain parameters occur, the numerical values of which have to be found from a comparison of the experimental and the theoretical absorption spectrum due to the corresponding transition. If certain small but systematic deviations between theory and experiment are neglected, values for the radiationless transition probability are found which are of a reasonable order of magnitude. It is furthermore shown in a qualitative way that the deviations may be explained by the fact that the theoretical spectrum is determined under the assumption that the Condon approximation is valid. For alkali-halide F-centres this is probably not allowed.

2224: B. H. Schultz: Surface recombination as a function of the concentration of charge carriers in the interior (Physica, **20**, 1031-1033, 1954).

Measured values of the surface recombination rate s of electrons and holes of P and N germanium are compared with what is to be expected from the considerations of Brattain and Bardeen on surface states. There is qualitative agreement: s is indeed large for materials of small resistivity, but it depends on the resistivity to a lesser extent than the theory predicts. For intrinsic germanium (i.e. germanium in which neither donors nor acceptors are dominant) which has been etch-polished, the value of s measured was 20 cm/sec.

2225: F. A. Kröger, H. J. Vink and J. Volger: Resistivity, Hall effect and thermo-electric power of conducting and photo-conducting single crystals of CdS, from 20—700 °K (Physica **20**, 1095-1099, 1954).

Short report on investigations which were described at length in Philips Res. Rep. **10**, 39-76, 1955: see these abstracts No. **R260**.

2226: G. H. Jonker: Semiconductor properties of mixed crystals with perovskite structure (Physica **20**, 1118-1122, 1954).

Mixed crystals of $La\,Me^{3+}O_3$ and $Sr\,Me^{4+}O_3$ (Me = Ti, Cr, Mn, Fe or Co) with perovskite structure have been prepared in ceramic form. The compounds with the general formula $[La_{1-x}Sr_x][Me^{3+}_{1-x}Me^{4+}_x]O_3$ show a high electric conductivity as the Me-ions are present in two valency states. Interesting properties are met in mixed crystals containing two kinds of metal ions of the transition group. Because of the difference in ionisation energies, one may expect that only one kind of the metals in these mixed crystals exists in two valency states. It is possible to prepare samples with high restivity, from which it can be concluded that one metal is present in the trivalent state and the other one in the tetravalent state, e.g. in $[La_{0.75}\,Sr_{0.25}][Fe^{3+}_{0.75}Mn^{4+}_{0.25}]O_3$. By the determination of the maxima of the resistivity in various series of mixed crystals, the following sequence of preference for the tetravalent state is found: Mn > Cr > Fe. Some of the measurements are complicated by the ferromagnetic properties of the materials, which have a strong influence on the conductivity. It was therefore difficult to find the place of Co in the sequence.

R 271: G. Klein: Rejection factor of difference amplifiers (Philips Res. Rep. **10**, 241-259, 1955, No. 4).

By analyzing the ordinary triode difference amplifier it is shown that the rejection factor can be made arbitrarily large, without the need for preselection of valves or stringent mutual equality of corresponding circuit elements. Some circuits are given which guarantee a high rejection factor, even with 10% difference of the corresponding components of the two halves. The theory is verified by a number of measurements.

R 272: J. Volger, J. M. Stevels and C. van Amerongen: Dielectric losses of various monocrystals of quartz at very low temperatures (Philips Res. Rep. **10**, 260-280, 1955, No. 4).

Between 14 and 150 °K the $\tan\delta$ vs T curve of quartz, measured at frequencies of 1 or 32 kc/s, may show a variety of maxima. The relaxation phenomena involved are correlated with both primary and radiation-induced lattice defects. Results of experiments with clear quartz, artificially irradiated quartz, natural smoky quartz and amethyst are reported. A discussion related to the nature of a number of lattice imperfections is given.

R 273: A. van Weel: Phase-linear television receivers (Philips Res. Rep. **10**, 281-298, 1955, No. 4).

A television receiver with phase-linear intermediate frequency amplifier is described. Normal selectivity demands are satisfied and conventional circuits are used. Performance compares very favouraby with receivers of which the I.F. phase errors are compensated for in the video-frequency part of either transmitter or receiver. The picture quality, apart from being optimum for the given bandwidth, is almost independent of tuning variations.